Structure and composition of bacterial communities in the rhizosphere and root of *Hypericum* species

Von der Fakultät für Lebenswissenschaften

der Technischen Universität Carolo-Wilhelmina zu Braunschweig

zur Erlangung des Grades einer

Doktorin der Naturwissenschaften

(Dr. rer. nat.)

genehmigte

D i s s e r t a t i o n

von Zahra Noviana

aus Bandung / Indonesia

1. Referent: Professor Dr. Jörg Overmann
2. Referent: Professor Dr. Michael Steinert
eingereicht am: 24.11.2020
mündliche Prüfung (Disputation) am: 10.02.2021

Druckjahr 2021

Vorveröffentlichungen der Dissertation

Teilergebnisse aus dieser Arbeit wurden mit Genehmigung der Fakultät für Lebenswissenschaften, vertreten durch den Mentor der Arbeit, in folgenden Beiträgen vorab veröffentlicht:

Publikation

Noviana, Z., Vieira, S., Pascual, J., Fobofou, S.A.T., Rohde, M., Spröer, C., Bunk, B. & Overmann, J. *Hypericibacter terrae* gen. nov., sp. nov. and *Hypericibacter adhaerens* sp. nov., two new members of the family *Rhodospirillaceae* isolated from the rhizosphere of *Hypericum perforatum*. International Journal of Systematic and Evolutionary Microbiology 70, 1850-1860 (2020).

Tagungsbeiträge

Noviana, Z., Sikorski, J. & Overmann, J., Structure and composition of bacterial communities in the rhizosphere of *Hypericum* species. Erstes PhytoAD-Symposium am IPB zu neuroaktiven Wirkstoffen aus Pflanzen, Halle (2019).

Posterbeiträge

Noviana, Z., Sikorski, J. & Overmann, J., Structure and composition of bacterial communities in the rhizosphere of *Hypericum* species (Poster) Vereinigung für Allgemeine und Angewandte Mikrobiologie Jahrestagung 2020, Leipzig (2020).

Dedicated to my beloved parents, Dr. H. Dede Margo Irianto and Dra. Hj. Yani Yusnia

Table of Contents

Chapter 1 – Summary

The plant genus *Hypericum* is known for its active constitutents, namely hypericin and hyperforin, with anti-anxiety, anti-depression, anti-inflammatory, anticancer, antiviral, and antimicrobial effects. Despite the fact that some bacterial isolates can trigger the production of hypericin and pseudo-hypericin in the seedlings and shoot cultures of *H. perforatum* when used as inoculant, the role of the entire rhizosphere and root bacterial community *in situ* has not yet been investigated. Soil is one of the largest bacterial diversity reservoirs on Earth, comprising up to a million bacterial phylotypes belonging to more than 50,000 species only in a single gram. Thus, the interaction between these diverse bacterial taxa along with biotic and abiotic factors may have an impact on the plant secondary metabolite profiles.

In the present thesis, culture-dependent and culture-independent techniques were employed to investigate the rhizosphere and root-endophyte bacterial communities of *Hypericum* plants, as a step closer in understanding the potential link between bacterial communities and the production of hypericin and hyperforin. Habitat type (either rhizosphere, inside the roots, or bulk soil) was determined as the strongest driver of bacterial communities of *Hypericum* plants, followed by soil substrate (with distinct pH) and plant species. Moreover, the impact of plant species was evident on the rhizosphere and root-endophyte bacterial communities but less pronounced on bulk soil communities. On the other hand, soil pH was the most important variable of bulk soil bacterial communities whereas the impact on the communities in the other two habitat types was evident, albeit weaker.

In agreement with what had been observed in previous studies, approximately 1% of bacterial taxa were active in bulk soil, the rhizosphere, and inside roots of *Hypericum* plants. Active taxa in the rhizosphere and roots of *Hypericum* include members of genera that are widely known as plant growth-promoting rhizobacteria (PGPR) such as *Rhizobacter, Bacillus,* and *Pseudomonas.* In addition, the genera *Gemmata, Haliangium, Pedomicrobium, Pirellula,* and *Planctomycetes* were identified as active but the ecological roles of these latter genera in association with plants is not sufficiently known so far. The specific active taxa in the rhizosphere and root of hypericin and hyperforin-producing species include members of *Streptomyces, Mesorhizobium, Bacillus, Bradyrhizobium*, and *Pseudonocardia*. Furthermore, two new members of the family *Rhodospirillaceae* were isolated from the rhizosphere of *H. perforatum* and further characterized based on taxonomic novelty. In addition, a member of *Acidobacteriota* that relies on helper bacteria to grow was retrieved from one of the soil sample.

This study provides novel insights regarding bacterial communities associated to *Hypericum* plants and forms the basis for more detailed analyses of the role of associated bacterial communities for plant secondary metabolite production. In particular, specific bacterial taxa could be identified that represent interesting targets for subsequent studies of hypericin and hyperforin production in *Hypericum*. Ultimately, the study provides a fundamental basis to understand the potential link between associated bacterial communities of *Hypericum* and hypericin and hyperforin content.

Chapter 2 – Introduction

2.1 – Plant and their prominent soil microbiome

The plant microbiome, also referred to the plant second genome, exerts beneficial impacts on plant health and productivity (Berendsen *et al.* 2012; Greenhalgh et al. 2016; Halfvarson et al. 2017). Since the very beginning of their evolution, plants have been living in a tight association with their microbiome. This long-term co-evolution has benefited both plant and the microbiome, including improvement of immune system on the plant side (Borrelli et al. 2018, Frantzeskakis et al. 2019). In the field of agriculture, the application of microbes can be more sustainable (Turner *et al.* 2013).

Some specific interactions between plant and their microbes, for example the formation of the root nodule in legume plant as a habitat for nitrogen-fixing bacteria (Maróti & Kondorosi 2014), are well studied. However, the full potential of the whole plant microbiome, especially those in the rhizosphere, to the plant health, productivity, and ecological functions remains mostly unexplored. This was also exacerbated by the fact that the majority of soil taxa still escape cultivation (Overmann 2013). Nevertheless, numerous culture-independent efforts, such as amplicon sequencing, metagenomic, metatranscriptomic, metaproteomic, and SIP-labelling, have been performed to discover specific interactions between the plant and their microbiome (Zarraonaindia 2013).

Plant microbiomes occur in the rhizosphere (the narrow area on the soil that is influenced by root secretions), endosphere (the area within the plant tissues), and phyllosphere (aboveground-surfaces of the plant). Each habitat provides different conditions for microbial life. The rhizosphere, first described as the plant-root interface (Hiltner, 1904), is the soil which is heavily influenced by plant root and rhizodeposition, including root exudation and mucilage secretion (Turner *et al.* 2013). While the rhizosphere is richer in nutrient content due to rhizodeposition effect (Nguyen 2003), the phyllosphere offers less resources and delivers more extreme environmental conditions (Vorholt 2012). Bacterial diversity in the rhizosphere (10^6 - 10^9 g^{-1}) is higher compared to phyllosphere (10^6 - 10^7 cm^{-2}) and root endosphere (10^4 - 10^8 g^{-1}) (Bulgarelli *et al.* 2013). In addition, the rhizoplane (root surface) is the area where the effects of root exudates and nutrient uptake are most pronounced (Wieland *et al.* 2001; Oh *et al.* 2012) (Figure 1).

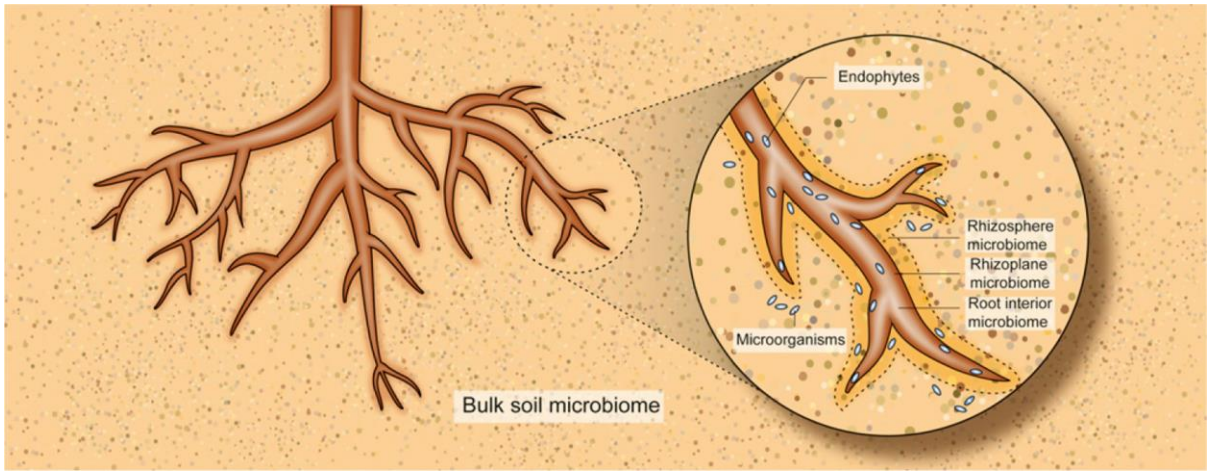


Figure 1 – Illustration of rhizosphere, rhizoplane, and root (endophytic) microbiomes (Gaiero *et al.* 2013).

2.1.1 – The diversity and significant importance of soil bacteria

In order to understand the ecological role of bacteria in the rhizosphere, we need to understand their distribution, functions, and interaction with other biotic and abiotic factors in the soil habitat. Soil represents a complex, heterogeneous, biodiversity-rich habitat where microorganisms, plants, and animals interact and carry important functions that are beneficial for human life (Wall *et al.* 2015). The complex interaction between these organisms and abiotic factors such as pH, nutrient content, oxygen availability, humidity, and temperature, is known to be important for the maintenance of soil ecosystem functions. The knowledge of how these soil organisms react to environmental changes is necessary to predict the impact of global warming or future environmental disturbances to the soil ecological functions.

Soil biodiversity is massive and represents one of the largest diversity reservoirs on Earth. Soil preserves the life of 300.000-400.000 plant species and thousands of earthworms, ants, and termites (Orgiazzi *et al.* 2016). Soil diversity is variable over time and is strongly affected by land management and climate change. The loss of taxa could affect soil ecosystem functions like the decomposition of organic matter or nutrient cycling. The high biological diversity of soil can provide functional redundancy which can act as a buffer to maintain soil ecosystem functions when specific taxa are lost, for example due to global warming (Wagg *et al.* 2019).

Soil fauna play an important role in carbon and nutrient cycling, soil aggregation, soil particle distribution, and in maintaining soil structures by the formation of earthworm cast or soil pores by radial pressures or ingestion by earthworm, termites, ants, etc. (Oades 1993; Osler & Sommerkorn 2007; Fry *et al.* 2012; Orgiazzi *et al.* 2016). Although bacteria and fungi contribute more than 1000 kg

microbial biomass per hectare of soil (Fierer 2017), their functional role in this habitat is still only partially understood (Maron *et al.* 2018).

The development of molecular methods has allowed the study of bacterial diversity in ever-increasing pace (Fierer *et al.* 2007). So far, the small subunit ribosomal RNA (SSU rRNA) and internal transcribed spacer (ITS) genes are mostly targeted for bacteria and fungi, respectively (Hug *et al.* 2016; Aslam *et al.* 2017). This has revealed that soil microorganism contributing one quarter of global microorganism diversity (at species level) on the planet (Wagg *et al*. 2019). In total, around 2 million fungal species and up to 10 million prokaryotic species are distributed globally (Wu *et al.* 2018). Of these, up to 10^{10} bacterial cells (Raynaud & Nunan 2014), which can correspond to 52,000 bacterial species (Roesch *et al.* 2010) or 10^6 bacterial phylotypes at most (Bickel & Or 2020), can occupy one gram of soil.

Microorganism contribute to more than 90% of the energy flow through a food chain in soil, including utilization of light energy by photoautotroph, inorganic compounds by chemoautotroph, and external organic carbon by heterotroph to produce energy sources such as adenosine triphosphate (ATP) (Adam & Duncan 2001; Morris & Blackwood 2015; van Elsas *et al.* 2019). Bacteria have been reported to play major roles in diverse soil biogeochemical processes, for example related to nitrogen, carbon, sulfur, and phosphorus cycle (Morris & Blackwood 2015; Lladó *et al.* 2017). The decomposition of dead plant and fungal biomass and the degradation of phenolic compounds (Lladó *et al.* 2017) are key function in the soil carbon cycle and chemical transformations of inorganic N, P, and S-compounds affect soil mineral availability. The role in nitrogen cycling is already well known. Unlike the majority of the living organism, some specific archaeal and bacterial taxa are capable of using dinitrogen gas (N_2) as their nitrogen source through nitrogen fixation. The resulting ammonia can then undergo nitrification process to first produce nitrite and then nitrate and can be converted back to N_2 via denitrification, all through bacterial-mediated step (Kuypers *et al.* 2018).

Soil microorganisms affect the climate through production or consumption of greenhouse gases, like carbon dioxide (CO_2), methane (CH_4), and nitrous oxide (N_2O). Microorganisms alter the atmospheric carbon dioxide concentration directly via heterotrophic respiration or indirectly by promotion of plant growth, leading to CO_2 uptake from the atmosphere. Also the balance of methane (CH_4) formation by methanogens and the utilization of methane by methanotrophs modulates the flux of methane between terrestrial and atmospheric environments. In addition, the flux of nitrous oxide (N_2O) highly depends on bacterial processes such as nitrification and denitrification (Orgiazzi *et al.* 2016). The role of soil as a source or a sink of these gases thus is governed by the activity of the resident taxa (Jansson & Hofmockel 2020).

Soil bacteria are also considered to be the important player in phosphorus cycle. Plant growth promoting rhizobacteria are known to produce organic acids and phosphatases/phytases as well, as the way to enhance the phosphate availability in the soil (Vacheron *et al.* 2013). In addition, plants depend on soil bacteria for their life survivability, such as for nitrogen and phosphorus uptakes, handling of environmental stresses, and protection against pathogen attack. Other important functions of soil bacteria include the production of siderophore to promote iron and trace elements uptake.

The application of soil microorganisms as biofertilizers and biocontrol agents can improve the chemical leaching or runoff, groundwater pollution, plant growth inhibition, etc. (Geisen *et al.* 2019). The response of microorganisms to specific environmental disturbances such as climate change have been intensively studied to predict the consequences of this disturbance and to find the solution to recover the soil ecological functions.

Microorganisms in soil are often limited by, and compete for nutrients, and sometime also oxygen (Fierer 2017). Only a small fraction of soil microorganisms is active (0.1 – 2%) while the rest are in a dormant state (Blagodatskaya 2013; Meyer *et al.* 2019). Nevertheless, soil is a dynamic habitat where changes in physicochemical conditions cause shifts in the composition of active taxa (Vieira *et al.* 2020). Specifically, the fraction of active microorganism can be 20 times higher in microbial hotspot such as the rhizosphere, detritusphere, and aggregate surfaces, where faster process rates and more intense interactions are observed compared to normal soil. In the rhizosphere, labile organics produced by plants and soil structures contribute to spatial and temporal heterogeneity of the microbial hotspot (Kuzyakov & Blagodatskaya 2015).

Inactive and low abundance bacterial taxa can act as a seed bank (Bickel & Or 2020) and can become active and increase in their abundance following environmental disturbance thereby maintaining soil ecosystem functions. Growth of rare taxa can also be promoted by favorable environmental conditions (Lennon & Jones 2011). The microbial states in soil can be categorised into active, potentially active, and dormant (Blagodatskaya 2013). The potentially active microbes can start their metabolism within minutes to few hours given appropriate substrates and suitable conditions, while it might take hours to days to resuscitate from a dormant state (Blagodatskaya 2013). In addition, Blazewicz *et al.* categorised microbial cell states as growing (actively dividing), active (metabolizing but not necessarily dividing), dormant (neither metabolizing nor dividing) and dead (Blazewicz *et al.* 2013). Those that are actively metabolizing (but not necessarily growing) are of immediate ecological relevance since they may be key players in some specific ecological functions.

One major drawback of previous studies focusing on microbial functioning is the failure to investigate active taxa since they mostly focused on the total biomass (Blagodatskaya 2013). This can potentially lead to ecologically wrong conclusions, as the compositions of the total fractions are often different from the active ones (Herzog *et al.* 2015; Romanowicz *et al.* 2016; Li *et al.* 2019). Thus, studies of microbial functioning should distinguish active taxa from the total microbial biomass. This can be achieved by utilization of various methods such as direct microscopy combined with cell staining, RNA-based fluorescent in situ hybridization (FISH) combined with the total microbial biomass staining, microbial growth-related approaches including substrate induced respiration (SIR) and substrate induced growth respiration (SIGR), and utilization of the rRNA:rDNA ratio (Blagodatskaya 2013; Loeppmann *et al.* 2018; Bowsher *et al.* 2019).

2.1.2 – The drivers of soil bacterial community

Despite the importance of soil bacteria to soil ecosystem services such as nutrient cycling, climate regulation, and biodiversity conservation, the functional, structural, and compositional responses of soil bacteria to environmental changes have not yet been fully understood (Orgiazzi *et al.* 2016). Soil bacterial communities exhibit remarkable stability against environmental disturbances due to their rapid growth, wide distribution, high metabolic flexibility, high tolerance to diverse environmental conditions, and their long-term evolutionary adaptation (Allison & Martiny 2008; Fuhrman *et al.* 2015; Jiao *et al.* 2019). The strategies of the soil bacteria to counterbalance environmental disturbances include resistance (tolerance to the disturbances), resilience (rapid recovery after the event of disturbances), and functional redundancy (preservation of ecological functions provided by different bacterial taxa with the same function (Allison and Martiny, 2008; Griffiths & Philippot 2013; Hodgson *et al.* 2015; Jiao et al., 2016).

According to the neutral and niche differentiation theory, stochastic and deterministic processes, respectively, may contribute to the assembly process of soil bacterial communities (Dumbrell *et al.* 2010; Ferrenberg *et al.* 2013). The latter theory emphasizes the resource preferences of the resident bacteria, occupying different niches (Leibold & McPeek 2006; Dumbrell *et al.* 2010), while the former suggests that inhabitants do not outcompete each other and therefore the communities are shaped based on stochastic process and dispersal limitation (Hubbel, 2001; Ferrenberg *et al.* 2013).

Several factors may drive deterministic community assembly, including soil edaphic properties or climate related-attributes, namely pH (Fierer & Jackson 2006, Landesman *et al*. 2014, Kaiser et al. 2016), organic carbon (Shen *et al.* 2015; Bickel & Or 2020), oxygen availability (Fierer 2017), moisture (Freedman and Zak 2015), nitrogen and phosphorus concentration (Shen *et al.* 2015; Dumbrell et al. 2010), texture and structure of the soil (Fierer 2017), temperature (Freedman and

Zak 2015), vegetation type (Knelman et al. 2012, Zhang et al. 2014, Mitchell et al. 2010), and predation and viral lysis effect, ordered from the more- to the less of important ones (Fierer 2017). Other factors affecting the communities include land use (Hartmann et al. 2014; Herzog et al. 2015; Berkelmann et al. 2018), elevation (Bryant et al. 2008, Singh *et al*. 2014), litter mass (Freedman and Zak 2015), and Fe and Ca content (Berkelmann et al. 2018).

Soil pH seems to be the best predictor of the community structure, diversity, and functions of bacterial communities in forests and temperate grasslands (Kaiser et al., 2016). However, the effect of pH may not be similar in different regions, as either neutral or alkaline pH have been reported to contribute to the highest measures of soil diversity in different regions (Fierer & Jackson 2006; Lauber *et al*. 2009; Kaiser et al., 2016). In addition, the response of soil bacteria to the changes of pH varies (Kaiser et al., 2016) and is not necessarily linked to their phylogenetic affiliation (Landesman et al., 2014). The relative abundance of some phyla such as Acidobacteriota, Verrucomicrobiota, and Proteobacteria has been reported to be correlated with the soil pH (Landesman et al., 2014; Shen *et al.* 2015). Acidophiles, such as *Acidithiobacillus* and *Sulfolobus*, include bacteria that tolerate extremely low pH, up to pH 1, and usually have pH optima of below pH 5.5. Alkalophiles prefer higher pH values, up to 11.5, and usually have pH optima of above 8.5. The survival of these bacteria at extreme pH is thought to be related to their cell wall composition (Kaiser et al., 2016) and their ability to maintain a constant value of their intracellular pH by applying organic and inorganic buffer system (phosphates, sulfonates, amines, alginic acid, etc.). Acidophiles maintain their intracellular pH at a value of 6 to 7 while alkaliphiles set theirs to be 1-2 pH points lower than the outside value. However, most prokaryotes are neutrophiles that have pH optima between 5 and 9, and keep their internal pH at a slightly alkaline condition (Overmann 2013).

The observation of bacterial communities along an elevational gradient in alpine tundra revealed an effect of elevation, along with the other important drivers such as total carbon, total nitrogen, C:N ratio, and dissolved organic carbon to the structure and diversity of the bacterial community (Shen *et al.* 2015). Land use has also been reported to have an impact on the soil bacterial community. High land-use intensity correlated with low diversity, as the most competitive species will become dominant in such a highly-stressed area. Conversely, higher diversity is observed at low land-use intensity (Tardy *et al.* 2017). In addition, the abundance of *Rhizobiales* was reported to decrease in fertilized soils, indicating that nitrogen fixation was not a selective trait (Berkelman *et al.* 2018).

Furthermore, investigation on 700 soil samples associated with 4 particular tree species, that were sampled from 12 different forests in the United States, revealed that different tree species host different bacterial community (Landesman *et al.* 2014). Bacterial communities across a long-term

chronosequence, which is a set of sites with similar ecological and edaphic attributes but varies in age, are affected by soil physical characteristics and temporal variation, with each accounted for 17% and 53% of the changes in phylogenetic structure of the community, respectively (Freedman & Zak 2015). The shift of abundance of members of *Alphaproteobacteria* and *Gammaproteobacteria* is mostly responsible for the compositional changes of bacterial communities along the chronosequence (Freedman & Zak 2015). Moreover, soil moisture was also reported to be a controlling variable and has a tight relation with richness and diversity of soil bacterial community (Angel *et al.* 2010; Yuan *et al*. 2014; Zhang *et al.* 2014; Shen *et al.* 2015).

Some of these factors exert combined effects on the bacterial communities (Rillig *et al.* 2019). For example, soil pH can indirectly alter soil nutrient content by modulating soil decomposition rate (Dumbrell *et al.* 2010) or nitrification processes (Zhalnina *et al.* 2014). In other cases, the effect of the tree species in altering soil microbial communities was due to their ability to modify soil pH (Dijkstra & Smits 2002; Reich *et al.* 2005; Landesman *et al.* 2014). However, only few studies have combined multiple factors of soil properties to test their importance in structuring microbial communities (Rillig *et al.* 2019).

Although some factors have been reported to be potential drivers of soil bacterial community, global drivers cannot be fully determined in many studies since they only covered specific regions of Earth. In addition, many studies failed to identify the non-linear relationship between potential drivers and bacterial community and only addressed a single or two parameters and thus cannot be used to adequately identify global predictors of bacterial community. For this reason, Delgado-Baquerizo and Eldridge investigated the bacterial communities from 237 locations belonging to arid, continental, and temperate ecosystems with ultraviolet light, climate (including aridity index, precipitation seasonality, maximum and minimum temperature, and the diurnal temperature range), soil physicochemical properties (including texture, pH, and the content of C, N, and P) and vegetation types (the presence or absence of forest and grassland) (Degado-Baquerizo & Eldridge 2019). The result demonstrates that only UV light, soil C, pH, and forest environment could be considered as global predictors of bacterial community in the three biomes (arid, continental and temperate climates) where the first three drivers are considered the most important ones. These four predictors are able to explain about 51 % variance of the bacterial communities across the biomes. In general, bacterial diversity peaked at intermediate UV light while soil C is negatively correlated with the diversity. The bacterial diversity increases with increasing pH, peaked at around pH 7-8 before decreasing at higher pH. Furthermore, the diversity is lower in forests than in grassland that is assumed to be related to the increasing soil C and reduced soil pH in the forest

environment. In such a rich and productive habitat, a strong competition will be promoted that may lead to the decrease of the bacterial diversity (Degado-Baquerizo & Eldridge 2019).

Moreover, a global meta-analysis of bacterial and archael communities has been performed to evaluate global predictors of the communities, which is included in the Earth Microbiome Project. According to the study, the bacterial and archaeal diversity can be differentiated into free-living or host-associated (Thompson *et al.* 2017). The former can be further classified into saline and non-saline environment and the latter into animal and plant association. Less richness is observed in the host-associated community when compared to the free living one, except for the rhizosphere richness that resembles soil community. Major differences in community composition is also observed between saline and non-saline environment. The predicted average community copy number (ACN) is higher in the animal-associated communities when compared to those of plant-associated or free-living ones, suggesting that higher ACN which allows maximum growth rates may be advantageous in the nutrient-rich habitat such as inside the animal tissues. Furthermore, the maximum bacterial and archaeal richness is observed at intermediate pH and temperature range (approximately at pH 7 and 19 $^{\circ}$C). The bacterial and archaeal diversity is also governed by nestedness instead of taxa turnover where the less diverse communities are more of a subset of the more diverse ones. In addition, the study also emphasizes the utilization of the individual 16S rRNA sequences (instead of clustering at 97% identity) to better capture habitat specificity of the bacterial and archael communities across the globe (Thompson *et al.* 2017).

2.1.3 – The crucial roles of rhizosphere bacterial community

Root exudation is one of the strongest drivers of the rhizosphere bacterial community. Root exudation profiles may vary depending on the plant species, cultivars, ages, and development stages, in addition to the contribution of environmental factors (Ling *et al.* 2013). In addition, the presence of plant pathogens can also affect the profile of the root exudates since they cause the release of specific antimicrobial compounds (Olanrewaju *et al.* 2019).

Root exudates mainly consist of organic acids and sugars, in addition to amino acids, fatty acids, growth factors, hormones, and antimicrobial compounds (Bertin *et al.* 2003) and affect the diversity and composition of the rhizosphere microbial communities (Badri & Vivanco 2009). Approximately 5 - 40% of photosynthetically-fixed carbon can be released to the rhizosphere as root exudates (Turner *et al.* 2013), making the rhizosphere a hotspot for microbial life. Beneficial microbes can be attracted by these compounds, which later play a role in promoting plant growth and productivity. Motility and chemotaxis, the production of lipopolysaccharides or biofilms, and fast growth rates are

reported to be important factors for root colonization (Lopes *et al.* 2016). The rhizosphere effect, which is the consequence of root exudation, drives rhizosphere community composition (Liu 2019). The rhizosphere community is characterized by lower diversity but significantly higher numbers of cells, larger cells sizes, and higher metabolic activities (Smalla *et al.* 2001; Dennis *et al.* 2010; Peiffer *et al.* 2013; Vieira *et al.* 2020).

Plant growth promoting rhizobacteria (PGPR) produce specific beneficial chemicals such as cytokinins, antioxidants, ACC deaminase, and volatile organic compounds that can assist the plant in mitigating the effect of drought, high salinity, and other environmental stresses (Figure 2) (Ryu 2009). As plant cannot directly utilize molecular nitrogen, PGPR provide the plant with accessible ammonia and nitrate through nitrogen fixation and nitrification (Donate-Correa 2004; Kamaruzzaman *et al.* 2020). In addition, PGPR are also capable of increasing phosphate availability in the soil by the production of either organic acid or phosphatases/phytases enzymes (Singh & Satyanarayana 2011).

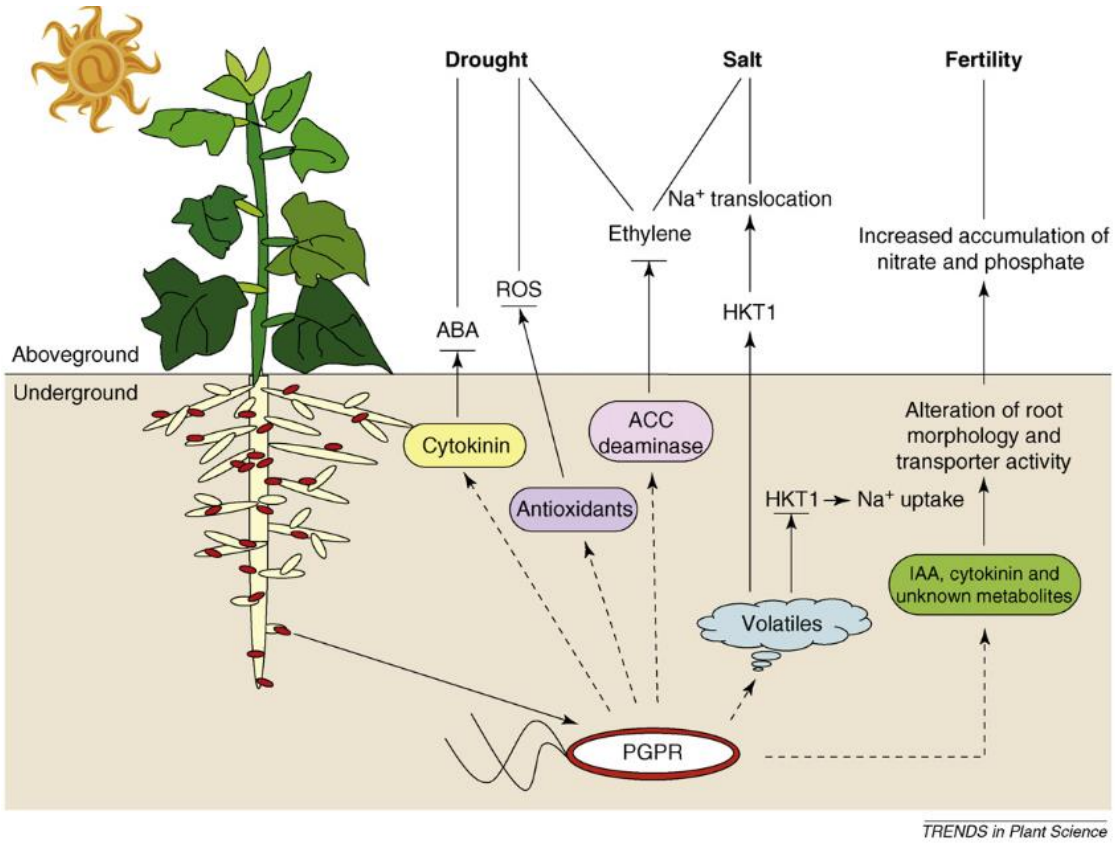


Figure 2 – Illustration of plant-growth-promoting rhizobacteria (PGPR) activities for improvement of the host plant health and productivity (Ryu 2009).

PGPR existence is also crucial to provide protection against plant pathogen, as they can have an antagonistic effect towards detrimental taxa by producing antimicrobial compounds. One example is

the production of diacetylphloroglucinol (DAPG) by *Pseudomonas fluorescens* which has an antifungal effect but also initiates transcription of specific genes of *Azospirillum brasiliense*, affecting root colonization and plant growth promotion (Combes-Meynet *et al.* 2011). In addition, some chemical signals such as flavonoids and strigolactones can elicit different responses from different taxa (Hassan & Mathesius 2012).

Moreover, PGPR can improve plant immunity by priming defense via systemic acquired resistance (SAR). In this process, the microbe-associated molecular patterns (MAMPs) of the PGPR, which include flagellin, elongation factor Tu, and chitin, are recognized by the plant and trigger plant immunity (Bittel & Robatzek 2007). Plants react by producing reactive oxygen species (ROS), deposition of callose, and activation of particular signaling and defense genes such as salicylic acid, jasmonic acid, and ethylene pathways. These pathways can then lead to an increase in the concentration of defensive enzymes and protein, along with secondary metabolites with protective properties. The promotion of plant growth will also be activated as a consequence of the pathways, and will further improve the defense mechanism against herbivorous insect (Figure 3).

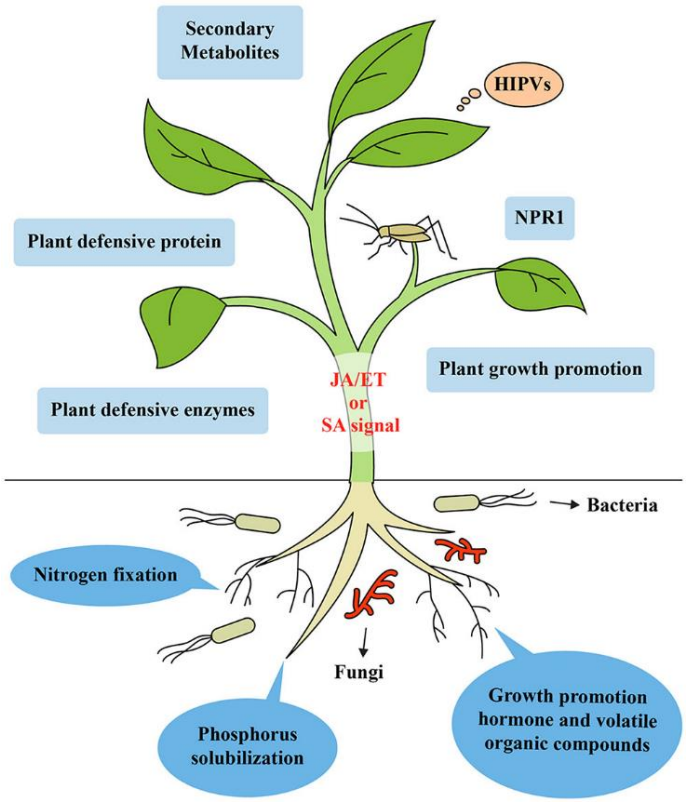


Figure 3 – The roles of beneficial microbes in promoting plant growth (Harun-Or-Rashid & Chung 2017).

2.1.4 – Roles of root endophytes for plant health and productivity

Endophytes, a term broadly used to define resident microbes inside the plant tissues, are thought to originate from soil (Compant *et al.* 2010). Soil bacteria often colonize the root through cracks or emerging lateral roots (Liu *et al.* 2017) (Figure 4). Fungi will release extracellular enzymes such as cellulose and pectin (Choi *et al.* 2013) as a strategy for initiating the colonization of roots.

Both culture-dependent and culture-independent approaches have revealed that endophyte diversity is lower compared to rhizosphere communities. However, some reports indicate that endophytes are more adapted and retain many functions related to the plant well-being. So far, the full potential of the endophytes could not be fully investigated, as many of the endophytes still resist cultivation efforts (Podolich 2015).

Endophytes are mainly non-pathogenic, but yet some are reported to be opportunistic (Sessitsch *et al.* 2012). In sugarcane and rice crops, culture-independent studies have revealed that *Alphaproteobacteria*, *Gammaproteobacteria*, and Firmicutes are among the common endophytes found inside the root (Fischer *et al.* 2012; Sessitsch *et al.* 2012). Furthermore, Actinobacteriota and some specific families of Proteobacteria were consistently found inside the root of *A. thaliana* (Lundberg et al. 2012; Bulgarelli *et al.* 2012).

The crucial roles of the root-endophytes have been reported to be similar to those of PGPR. These include nitrogen fixation (James 2000), phytohormones and siderophore production, phosphate solubilization, and 1-aminocyclopropane 1-carboxylic acid deaminase activities (Compant *et al.* 2010; Sessitsch *et al.* 2012; Suarez-Moreno *et al.* 2012). The latter is related to the protection against environmental stresses by inhibiting plant ethylene pathway (Glick 2005). In addition, endophyte is also known to produce some protective molecules, such as melatonine, proline, and carotenoids (Pacifico *et al.* 2019).

Endophytes can also provide protection against phytopathogenic microbes, either by antagonistic effects (Monteiro *et al.* 2012) or by seizing the ecological niches that can be occupied by the plant pathogens. On the basis of this antagonistic relation, utilization of endophytes as biocontrol agents has been employed to replace and minimize negative effects of chemical pesticide (Pacifico *et al.* 2019). One study suggests that, of several endophyte isolates, genera such as *Bacillus* and *Pantoea* are the most effective agents to induce resistance against fungal *Botrytis cinerea* in grapevine plant (Campisano *et al.* 2015). In addition, *Pseudomonas fluorescens, Bacillus subtilis,* and *Malus domestica* can be utilized as biocontrol agents against crown gall diseased caused by *Rhizobium vitis* (Eastwell *et al.* 2006).

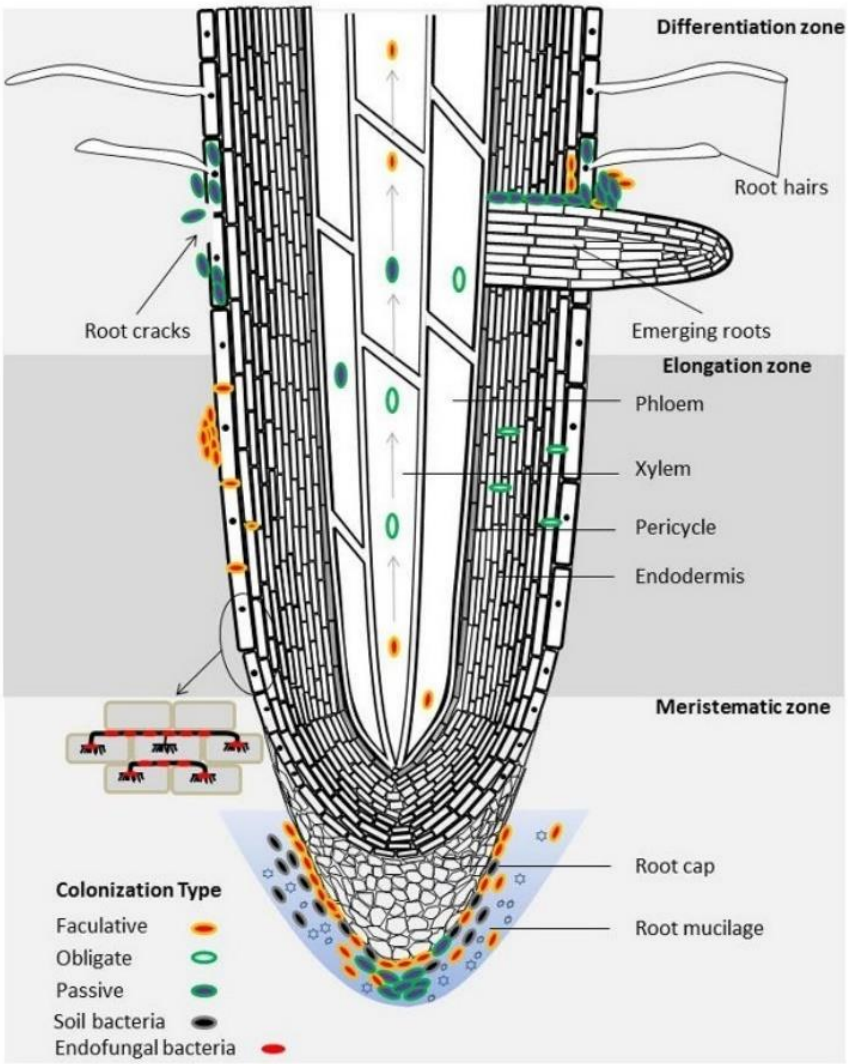


Figure 4 – Illustration of root habitat where soil bacteria can enter through cracks and emerging lateral root (Liu *et al.* 2017).

2.2 – The plant genus *Hypericum* with active compounds to mitigate anxiety and depression

Hypericum is a highly diverse and widely distributed plant genus found in tropical, subtropical, and temperate regions, and comprising approximately 500 species that were assigned to 36 different sections (taxonomic rank between genus and species used in botany) (Nürk *et al.* 2013). The genus is very well-known for the production of hyperforin and hypericin with anti-anxiety and anti-depression effect (Figure 5). The hypericin and hyperforin producing-plant species is usually observed to retain two glandular types, the dark glands and translucent glands (Figure 6), producing hypericin and hyperforin, respectively. The most notable species is *Hypericum perforatum*, in which the extract is commercially available as herbal medicine to treat against depression and anxiety (Kusari *et al.* 2009).

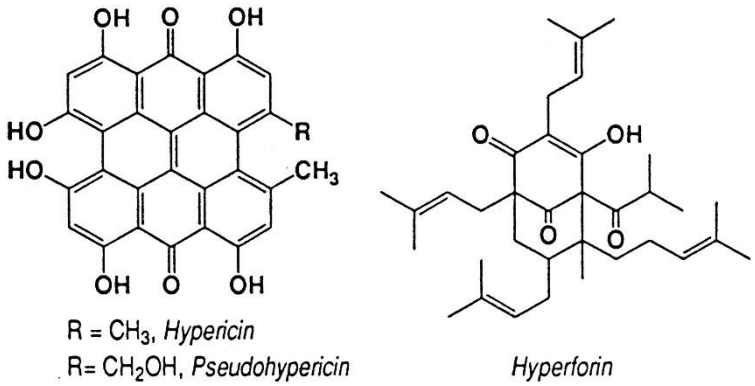


Figure 5 – Chemical structure of hypericin, pseudohypericin, and hyperforin (Briskin 2000).

It has been reported that approximately 60% of the *Hypericum* sections is able to synthesize hypericin (Košuth *et al.* 2011). The mode of action of hypericin in mitigating anxiety and depression symptoms is thought to be linked to the inhibition of monocyte cytokine production of interleukin-6 while hyperforin mode of action might be related to the inhibition of serotonin, norepinephrine, and acetylcholine uptake (Kinrys *et al.* 2009). Hypericin has also been reported to have antiproliferative, anti-inflammatory, antiviral, and antimicrobial properties (Cirak *et al.* 2016; Wang *et al.* 2019). Also, antibacterial and antiviral properties were reported for hyperforin. Methicillin-resistant (MRSA) and penicillin-resistant (PRSA) *Staphylococcus aureus* were especially susceptible to this compound (Klemow *et al.* 2011).

Figure 6 – Dark (black arrow) and translucent (white arrow) glands can be seen on the leaves of *H. perforatum* plant that was cultivated in the greenhouse of the Botanic Garden and Botanical Museum (BGBM) Berlin.

Previous study suggested that the hypericin and pseudo-hypericin content is highest in *H. montanum* while hyperforin and quercetin content is highest in *H. perforatum* (Kusari *et al.* 2009). Other hypericin and hyperforin producers include *H. hirsutum* and *H. maculatum* (Kusari *et al.* 2009; Cirak *et al.* 2016) while *H. polyphyllum* produce hypericin (Stojanović 2013) and hyperforin (Raclariu *et al.* 2017). *H. olympicum* has been reported as either producer of the compounds (Kitanov 2001; Smelcerovic 2006; Stojanović 2013) or a non-producer (Crockett *et al.* 2005). This inconsistency was most likely due to differences in the detection limit of the different studies (Kusari *et al.* 2015; Cirak *et al*. 2016). Moreover, the two compounds were not detected in *H. androsaemum* (Kitanov 2001; Aziz *et al.* 2006; Caprioli 2016) and *H. balearicum* (Stojanović 2013).

As had been mentioned above, the plant secondary metabolites can be influenced by beneficial microbes in the root via jasmonic acid, ethylene, or salicylic acid pathways (Harun-Or-Rashid & Chung 2017). Correspondingly, the hypericin content of *H. perforatum* can be induced by the exposure of the plant cell cultures to jasmonic acid in the dark (Walker *et al.* 2001). However, the study concerning the impact of the associated bacterial community of *Hypericum* on hypericin and hyperforin production is very limited. It is only known so far that the hypericin and pseudo-hypericin content can be changed by inoculation with specific bacteria, such as *Pseudomonas fluorescens* and *Stenotrophomonas maltophilia,* to the seedling of *H. perforatum in vitro* (Mañero *et al.* 2012). In order to gain more insights regarding the link of the associated bacterial community and hypericin and hyperforin, a full investigation of the bacterial community in the natural habitat should be performed.

2.3 – Approaches towards the analysis of the diversity of soil bacteria

Culture-independent techniques such as metagenomics have revealed an enormous microbial diversity in soils. Two common molecular approaches are the random-shotgun sequencing, which involves sequencing of all genomic fragments in the sample, and a targeted sequencing after amplification of only specific marker genes (Fuhrman 2012). 16S ribosomal RNA gene sequences have widely been used as a marker gene for bacterial community analysis (Tringe & Hugenholtz 2008; Chaudhary *et al.* 2015) but these analyses are limited by the short-read lengths (Delmont et al. 2015). State-of-the-art technology, such as third generation sequencing, enables direct sequencing of >10,000 bases in the one read, which is 10 times exeeding read lengths of Sanger sequencing (McCarthy 2010).

The culture-independent approach has revealed the vast majority of soil bacteria to be not yet culturable (Figure 7), due to the lack of information and proper techniques to cultivate them in the

laboratory (Stewart 2012). The fact that a significant number of bacterial taxa are unculturable was first recognized as "The great plant count anomaly" where the number of cells observed under the microscope outweighed the number of colonies grow in the cultivation agar media (Connon *et al.* 2002) by several orders of magnitude (Staley & Konopka 1985; Stewart 2012).

The fraction of the culturable taxa is predicted to be only between 0.1% and 0.001%, while the rest of the taxa still resist cultivation (Pham & Kim 2012; Overmann 2013). In order to better understand the metabolic capabilities and ecological roles of these bacteria, axenic cultivation and characterization of metabolic pathways is usually necessary (Stewart 2012; Overmann 2013). Cultivation is also important to retrieve many taxa that have been reported to be active within soil habitat (Stewart 2012), for example the genus *Nitrospira* that only has 9 representative isolates at the time of the writing, albeit recognized as one of the key players in nitrogen cycling (Daims *et al.* 2015), along with those major players in carbon, nitrogen, and other elements cycling, novel natural products producer, and the ones that have an impact to their neighboring organisms or other environmental processes (Stewart 2012).

Cultivation is strongly biased towards just 4 bacterial phyla, as approximately 90% of isolated species are classified as members of Proteobacteria, Firmicutes, Actinobacteriota, and Bacteroidota (Overmann *et al.* 2017; Chaudhary *et al.* 2019). It also has been reported that, out of approximately 80 recognized phyla, only 30 have isolated representatives. In contrast, relatively low numbers of representative isolates are available for some of the abundant soil phyla, such as the *Acidobacteriota* (Sait *et al.* 2002; Fierer *et al.* 2007), which produce diverse enzymes for the hydrolysis of cellulose, starch, and chitin, that are important for soil ecological functions (Lladó 2016). Moreover, novel strains often exhibit a high 16S rRNA gene sequence similarity to the isolated taxa, indicating a close phylogenetic relation (Overmann 2013). Since phylogenetically close bacterial taxa often occupy similar niches and perform the same function (Li *et al.* 2019), cultivation should target distant relatives of bacteria that has been cultivated so far to reveal more insights regarding ecological roles of soil bacteria.

2.3.1 – Strategies to improve culture-independent approaches

Microbial community analysis based on culture-independent techniques harbour several limitations, including the taxonomic resolution. In order to reveal more about potential roles of soil bacteria, the study needs to be performed at a deeper taxonomic level, even better at a cell-specific level. Recent study suggests that even community analysis based on the full length of 16S rRNA amplicon sequencing (~1500 bp) cannot fully provide taxonomic resolution at species and strain level when clustered below 99% similarity, since it cannot resolve nucleotide insertions/deletions that occur

between the 16S rRNA copies (Johnson *et al.* 2019). Therefore, it is better to analyse each of individual sequence (sequence variant level) or perform sequencing of all microbial genomes in the sample (metagenomics). However, metagenomic is only appropriate to analyse the most abundance taxa in a microbial-rich habitat like soil due to the problem regarding the sample coverage. For a large dataset comprising hundred of soil samples, amplicon sequencing is still preferable to access microbial diversity compared to metagenomics since it is more practical (metagenomic assembly is currently a computationally demanding process) and provides a better coverage of microbial diversity (Quince *et al.* 2017).

Moreover, community analysis based on amplicon sequencing and metagenomic is lacking due to the failure in identifying active taxa and thus cannot address functional capacity of the microbial community. In order to investigate the active fraction of microbial community, sequencing can be coupled with methods such as direct staining of nucleic acids and cell wall components and RNA-based FISH (Blagodatskaya 2013). Moreover, DNA-, protein-, carbohydrate-, and phenotype arrays, as well quantitative PCR can be performed to gain more insights regarding the functional role of soil microbial community (Zhou *et al.* 2015).

Furthermore, analysis in a more targeted manner can also be employed. Cell or taxa of interest can be deeply investigated using flow-sorting or single-cell genomics. Metagenomics can be combined with *in situ* enrichment for deeper analysis of the functional role of specific members of the community. In addition, immunomagnetic separation using antibody can isolate specific cells that can be further isolated or investigated by sequencing (Quince *et al.* 2017).

2.3.2 – Strategies to improve cultivation success

Several factors contribute to the low success of cultivation, and are often linked to the failure to replicate growth conditions of target bacteria like the Candidate phylum TM7 in their natural habitat (Stewart 2012). Other limiting factors include the physiological state of target bacteria, employment of nutrient-rich growth media that only targeted copiotroph but not the oligotrophs, lack of consideration of bacterial interactions, and short incubation period that results in retrieval of only fast-growing taxa (Overmann *et al.* 2017). In addition, the unknown adaptations of the target bacteria, such as their substrate preferences, appropriate temperature and pH for growth, osmotic condition, oxygen availability, and multiplication period (Stewart 2012; Chaudhary *et al.* 2019), make the cultivation effort more challenging.

Due to the fact that some nutrients are limited in soil, many soil taxa were found to be in their dormant state as a strategy to increase their survival (Overmann et al. 2017). The dormant state is

reported to be a pitfall for cultivation effort, as the bacteria in this state often escape cultivation, especially with the nutrient-rich media being used as the growth media. Therefore, utilization of growth media with low nutrient content is advisable (Azevedo et al. 2012).

Other factors that were believed to limit cultivation are the absence of essential growth factors or signaling molecules. The addition of the resuscitation-promoting-factors (Rpf) has been reported to be crucial in reviving *Micrococcus luteus* from its dormant state and also beneficial to promote growth of *Mycobacterium tuberculosis* and other members of Actinobacteriota (Kell & Young 2000; Vartoukian *et al.* 2010). It is also believed that due to the absence of appropriate growth factors and signaling molecules, bacteria could enter a very low metabolic activity and thus the incubation period needs to be extended further to successfully retrieve them (Vartoukian *et al.* 2010).

One of the strategies to improve cultivation success is to try to provide the target bacteria with the nutrients available in their natural habitat (Chaudhary *et al.* 2019). With this idea in mind, Epstein and Lewis had designed a semipermeable diffusion chamber where the target bacteria can be cultured with additional nutrients from their original habitat (such as addition of seawater to the cultivation media) without the risk of contamination (Kaeberlein *et al.* 2002; Stewart 2012). The utilization of seawater as growth media, along with the employment of dilution to extinction technique, was successful to retrieve members of the abundant marine SAR11 clade, that has been predicted to play major roles in carbon and energy balance in their natural habitat (Rappé *et al.* 2002).

The dilution to extinction technique, where the number of cells inoculated to the growth media were intentionally limited up to the point that only one or a few bacteria are present in inoculum, is also advantageous in the prevention of competition between the target bacteria and the fast-growing ones or to avoid any taxa that produce antimicrobial compounds (Puspita *et al.* 2012; Stewart *et al.* 2012).

Furthermore, a high number of soil bacteria are considered to be oligotrophic prefering nutrient-poor over nutrient-rich media for their growth. Therefore, many recent cultivation approaches were shifted toward utilization of growth media with low-nutrient content. In addition, the growth rate of oligotrophs is often significantly lower than that of fast-growing bacteria, and thus longer incubation period needs to be introduced (Puspita *et al.* 2012; Lladó 2016).

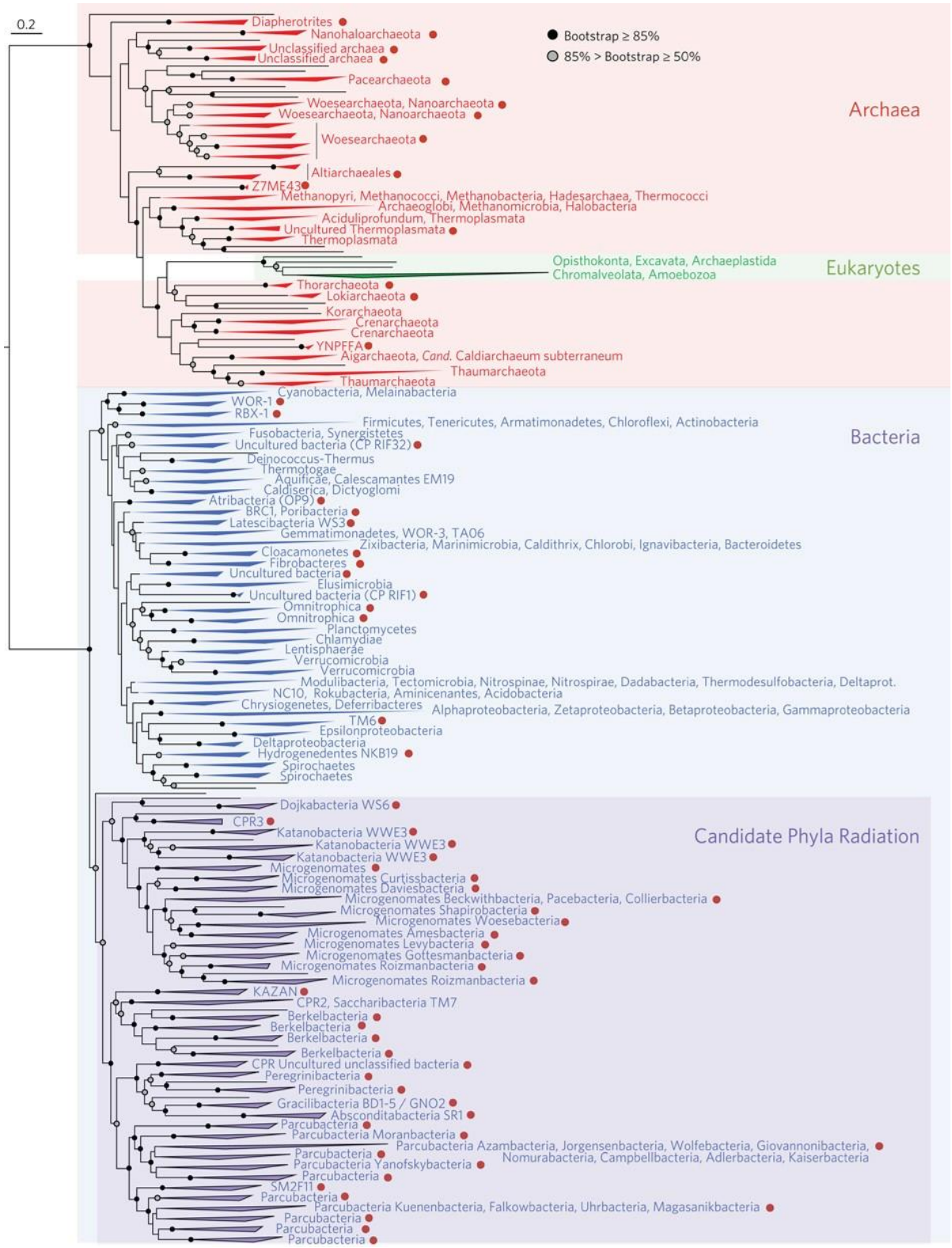


Figure 7 – The first comprehensive tree of life based on molecular information of all three domains (Archaea, Eukaryotes, and Bacteria). Red dots depict lineages with no representative isolates, including members of Candidate Phyla Radiation belonging to Bacteria domain (Hug *et al.* 2016).

Additionally, the co-cultivation approach can be introduced to improve cultivation success. Some bacterial taxa have been observed to grow only with other bacteria in close proximity, requiring growth in coculture (Kaeberlein *et al.* 2002; Stewart 2012). Some of the lineages, especially members of the so-called "Candidate Phyla Radiation" have limited metabolic capabilities that are most likely due to their relatively small genome size compared to other phyla (Hug *et al.* 2016). Thus, the cultivation effort can be adjusted by introducing helper bacteria with the ability to produce crucial compounds that are essential for the growth of the target taxa (Chaudhary *et al.* 2019). The addition of supernatant of target bacteria has also been reported to improve cultivation success, as demonstrated with the increase of growth of the previously-uncultured *Catellibacterium nectariphilum* in the presence of supernatant of another bacterium (Tanaka *et al.* 2004). Moreover, high throughput RNA sequencing technique (RNA-seq) can be utilized to determine the highly expressed genes of the target bacteria in their natural habitat and the information can be used to design the appropriate growth media, which has led to the isolation of the murin-dependent *Rikenella*-like bacterium resided in the gut of leech (Bomar *et al.* 2011).

A targeted cultivation considering the time of the sampling, oligotrophic media, longer incubation periods, addition of growth factors and signaling molecules, and a small inoculum size, has been proven to increase the success-rate of cultivation compared to conventional methods (Chaudhary *et al.* 2019). In addition, some of the novel taxa have been successfully retrieved with cocultivation with *Escherichia coli* or with amoeba (Overmann 2013). A high throughput cultivation with the utilization of all mentioned information will, without a doubt, increase the rate of the discovery of novel taxa, which needs to be done in order to unravel the full potential of soil taxa in regard to the soil ecological functions.

2.4 – Aims of the study

As a step forward in understanding the link between bacterial community and hypericin and hyperforin production in *Hypericum* plants, the rhizosphere and root-associated bacterial communities were investigated by targeting the V3 region of 16S rRNA gene sequences. The impact of habitat type (bulk soil, the rhizosphere, and the inner root), soil substrate (with distinct pH), and plant species on *Hypericum* bacterial communities were determined. Both total (DNA-based metabarcoding) and active (cDNA libraries were constructed from RNA prior to sequencing) bacterial communities were investigated for each habitat. The active bacterial taxa were determined based on rRNA:rDNA ratio values. The high throughput cultivation in liquid media was performed to retrieve phylogenetically novel bacterial taxa associated with *Hypericum* which might modulate hypericin and hyperforin. This cultivation effort is also intended to access soil diversity where most

of the taxa still refuse to be axenically cultured in the laboratory and to explore less-studied functions of novel bacterial taxa. Strategies that were employed for the cultivation include a targeted Illumina sequencing, utilization of nutrient-poor media, long incubation period and low number of inoculum size. The novel isolates will be further tested and fully characterized in the laboratory, for their chemotaxonomic and genomic properties.

2.5 – References

Adam, G. & Duncan, H. 2001. Development of a sensitive and rapid method for the measurement of total microbial activity using fluorescein diacetate (FDA) in a range of soils. *Soil Biology and Biochemistry*, **33**, 943-951, doi: 10.1016/S0038-0717(00)00244-3.

Allison, S. D., & Martiny, J. B. H. 2008. Resistance, resilience, and redundancy in microbial communities. *Proceedings of the National Academy of Sciences*, **105**, 11512-11519, doi: 10.1073/pnas.0801925105.

Angel, R., Soares, M. I. M., Ungar, E. D. & Gillor, O. 2010. Biogeography of soil archaea and bacteria along a steep precipitation gradient. *The ISME Journal*, **4**, 553-563, doi: 10.1038/ismej.2009.136.

Aslam, S., Tahir, A., Aslam, M. F., Alam, M. W., Shedayi, A. A. & Sadia, S. 2017. Recent advances in molecular techniques for the identification of phytopathogenic fungi – a mini review. *Journal of Plant Interactions*, **12**, 493-504, doi: 10.1080/17429145.2017.1397205.

Azevedo, N. F., Bragança, S. M., Simes, L. C., Cerqueira, L., Almeida, C., Keevil, C. W. & Vieira, M. J. 2012. Proposal for a method to estimate nutrient shock effects in bacteria. *BMC Research Notes*, **5**, doi: 10.1186/1756-0500-5-422.

Aziz, N., Sauve, R. J., Long, D. & Cherry, M. 2006. Genetic and Phytochemical Diversity Assessment Among Eleven *Hypericum* Accessions via AFLP and HPLC Analyses. *Journal of Herbs, Spices & Medicinal Plants*, **12**, 97–105, doi: 10.1300/J044v12n01_09.

Badri, D. V. & Vivanco, J. M. 2009. Regulation and function of root exudates. *Plant, cell & environment*, 666-681, doi: 10.1111/j.1365-3040.2009.01926.x.

Badri, D. V., Zolla, G., Bakker, M. G., Manter, D. K. & Vivanco, J.M. 2013. Potential impact of soil microbiomes on the leaf metabolome and on herbivore feeding behavior. *New Phytologist*, **198**, 264-273, doi: 10.1111/nph.12124.

Berendsen, R., Pieterse, C. M. J & Bakker, P. A. H. M. 2012. The rhizosphere microbiome and plant health. *Trends in Plant Science*, **8**, 478-486, doi: 10.1016/j.tplants.2012.04.001.

Berkelmann, D., Schneider, D., Engelhaupt, M., Heinemann, M., Christel, S., Wijayanti, M., Meryandini, A. & Daniel, R. 2018. How Rainforest Conversion to Agricultural Systems in Sumatra (Indonesia) Affects Active Soil Bacterial Communities. *Frontiers in Microbiology*, **9**, doi: 10.3389/fmicb.2018.02381.

Bertin, C. Yang, X. H. & Weston, L. A. 2003. The role of root exudates and allelochemicals in the rhizosphere. *Plant and Soil*, **256**, 67-83, doi: 10.1023/A:1026290508166.

Bickel, S. & Or, D. 2020. Soil bacterial diversity mediated by microscale aqueous-phase processes across biomes. *Nature Communications*, **11**, doi: 10.1038/s41467-019-13966-w.

Bittel, P. & Robatzek, S. 2007. Microbe-associated molecular patterns (MAMPs) probe plant immunity. *Current Opinion in Plant Biology*, **10**, 335-341, doi: 10.1016/j.pbi.2007.04.021.

Blagodatsky, S., Heinemeyer, O. & Richter, J. 2000. Estimating the active and total soil microbial biomass by kinetic respiration analysis. *Biology and Fertility of Soils*, **32**, 73–81, doi: 10.1007/s003740000219.

Blagodatskaya, E. & Kuzyakov, Y. 2013. Active microorganisms in soil: Critical review of estimation criteria and approaches. *Soil biology and biochemistry*, **67**, 192-211, doi: 10.1016/j.soilbio.2013.08.024.

Blazewicz, S. J., Barnard, R. L., Daly, R. A., & Firestone, M. K. 2013. Evaluating rRNA as an indicator of microbial activity in environmental communities: limitations and uses. *The ISME Journal*, **7**, 2061-2068, doi: 10.1038/ismej.2013.102.

Bomar, L., Maltz, M., Colston, S. & Graf, J. 2011. Directed culturing of microorganisms using metatranscriptomics. *MBio*, **2**, doi: 10.1128/mBio.00012-11.

Borrelli, G. M., Mazzucotelli, E., Marone, D., Crosatti, C., Michelotti, V., Valè, G., & Mastrangelo, A. M. 2018. Regulation and evolution of NLR genes: a close interconnection for plant immunity. *International Journal of Molecular Sciences*, **19**, doi:10.3390/ijms19061662.

Bowsher, A. W., Kearns, P. J. & Shade, A. 2019. The 16S rRNA/rRNA Gene Ratios and Cell Activity Staining Reveal Consistent Patterns of Microbial Activity in Plant-Associated Soil. *mSystems*, **4**, doi: 10.1128/mSystems.00003-19.

Briskin, D. P. 2000. Medicinal Plants and Phytomedicines. Linking Plant Biochemistry and Physiology to Human Health. *Plant Physiology*, **124**, doi: 10.1104/pp.124.2.507.

Brown, G. G., Barois, I. & Lavelle, P. 2000. Regulation of soil organic matter dynamics and microbial activity in the drilosphere and the role of interactions with other edaphic functional domains. *European Journal of Soil Biology*, **36**, 177-198, doi: 10.1016/S1164-5563(00)01062-1.

Bryant, J. A., Lamanna, C., Morlon, H., Kerkhoff, A. J., Enquist, B. J. & Green, J. L. 2008. Microbes on mountain sides: contrasting elevational patterns of bacterial and plant diversity. *Proceedings of the National Academy of* Sciences, 105, 11505–11511, doi: 10.1073/pnas.0801920105.

Bulgarelli, D., Rott, M., Schlaeppi, K., Ver Loren van Themaat, E., Ahmadinejad, N., Assenza, F., Rauf, P., Huettel, B., Reinhardt, R., Schmelzer, E., Peplies, J., Gloeckner, F. O., Amann, R., Eickhorst, T. & Schulze-Lefert, P. 2012. Revealing structure and assembly cues for Arabidopsis root-inhabiting bacterial microbiota. *Nature*, **488**, 91-95, doi: 10.1038/nature11336.

Bulgarelli, D., Schlaeppi, K., Spaepen, S., van Themaat, E. V. L. & Schulze-Lefert, P. 2013. Structure and functions of the bacterial microbiota of plants. *Annual Review of Plant Biology*, **64**, 807–838, doi: 10.1146/annurev-arplant-050312-120106.

Campisano, A., Pancher, M., Puopolo, G., Puddu, A., Lopez-Fernandez, S., Biagini, B., Yousaf, S., & Pertot, I. 2015. Diversity in endophyte populations reveals functional and taxonomic diversity between wild and domesticated grapevines. *The American Journal of Enology and Viticulture*, **66**, 12 – 21. doi: 10.5344/ajev.2014.14046.

Caprioli, G., Alunno, A., Beghelli, D., Bianco, A., Bramucci, M., Frezza, C., Iannarelli, R., Papa, F., Quassinti, L., Sagratini, G., Tirillini, B., Venditti, A., Vittori, S. & Maggi, F. 2016. Polar Constituents and Biological Activity of the Berry-Like Fruits from *Hypericum androsaemum* L. *Frontiers in Plant Science*, **7**, doi: 10.3389/fpls.2016.00232.

Chaudhary, D. K., Khulan, A. & Kim, J. 2019. Development of a novel cultivation technique for uncultured soil bacteria. *Scientific Reports*, **9**, doi: 10.1038/s41598-019-43182-x.

Chaudhary, N., Sharma, A. K., Agarwal, P., Gupta, A. & Sharma, V. K. 2015. 16S Classifier: A Tool for Fast and Accurate Taxonomic Classification of 16S rRNA Hypervariable Regions in Metagenomic Datasets. *PLoS* ONE, **10**, doi: 10.1371/journal.pone.0116106.

Choi, J. Kim, K-T., Jeon, J. & Lee, Y-H. Fungal plant cell wall-degrading enzyme database: a platform for comparative and evolutionary genomics in fungi and Oomycetes. *BMC Genomics*, **14**, doi: 10.1186/1471-2164-14-S5-S7.

Cirak, C., Radusiene, J., Jakstas, V., Ivanauskas, L., Seyis, F. & Yayla, F. 2016. Secondary metabolites of seven *Hypericum* species growing in Turkey. *Pharmaceutical Biology*, **54**, doi: 10.3109/13880209.2016.1152277.

Combes-Meynet, E., Pothier, J. F., Moenne-Loccoz, Y. and Prigent-Combaret, C. 2011. The Pseudomonas secondary metabolite 2,4-diacetylphloroglucinol is a signal inducing rhizoplane expression of Azospirillum genes involved in plant growth promotion. *Molecular Plant-Microbe Interactions*, **24**, 271-284, doi: :10.1094 / MPMI -07-10-0148.

Compant, S., Clement, C., Sessitsch, A. 2010. Plant growth-promoting bacteria in the rhizo- and endosphere of plants: their role, colonization, mechanisms involved and prospects for utilization. *Soil Biology and* Biochemistry, **42**, 669-678, doi: 10.1016/j.soilbio.2009.11.024.

Connon, S. A. & Giovannoni S. J. 2002. High-Throughput Methods for Culturing Microorganisms in Very-Low-Nutrient Media Yield Diverse New Marine Isolates. *Applied and Environmental Microbiology*, **68**, 3878-3885, doi: 10.1128/AEM.68.8.3878-3885.2002.

Crockett, S. L., Schaneberg, B. & Khan, I. A. 2005. Phytochemical profiling of new and old world *Hypericum* (St. John's Wort) species. *Phytochemical Analysis*, **16**, 479-485, doi: 10.1002/pca.875.

Daims, H., Lebedeva, E. V., Pjevac, P., Han, P., Herbold, C., Albertsen, M., Jehmlich, N., Palatinszky, M., Vierheilig, J., Bulaev, A., Kirkegaard, R. H., von Bergen, M., Rattei, T., Bendinger, B., Nielsen, P. H. & Wagner, M. 2015. Complete nitrification by *Nitrospira* bacteria. *Nature,* **528,** 504–509, doi: https://doi.org/10.1038/nature16461.

Delgado-Baquerizo, M. & Eldridge, D. J. 2019. Cross-Biome drivers of soil bacterial alpha diversity on a worldwide scale. *Ecosystems*, **22**, 1220-1231, doi: 10.1007/s10021-018-0333-2.

Delmont, T. O., Eren, A. M., Maccario, L., Prestat, E., Esen, Ö. C., Pelletier, E., Le Paslier, D., Simonet, P. & Vogel, T. M. 2015. Reconstructing rare soil microbial genomes using in situ enrichments and metagenomics. *Frontiers in Microbiology*, **6**, doi: 10.3389/fmicb.2015.00358.

Dennis, P. G., Miller, A. J. & Hirsch, P. R. 2010. Are root exudates more important than other sources of rhizodeposits in structuring rhizosphere bacterial communities? *FEMS Microbiology Ecology*, **72**, 313-327, doi: 10.1111/j.1574-6941.2010.00860.x.

Dijkstra, F. & Smits, M. 2002. Tree Species Effects on Calcium Cycling: The Role of Calcium Uptake in Deep Soils. *Ecosystems*, **5**, 385–398, doi: 10.1007/s10021-001-0082-4.

Donate-Correa, J., Le´on-Barrios, M. & Pere´z-Galdona, R. 2004. Screening for plant growth-promoting rhizobacteria in *Chamaecytisus proliferus* (tagasaste), a forage tree-shrub legume endemic to the Canary Islands. *Plant and Soil*, **266**, doi: 10.1007/s11104-005-0754-5.

Dumbrell, A. J., Nelson, M., Helgason, T., Dytham, C. & Fitter, A. H. 2010. Relative roles of niche and neutral processes in structuring a soil microbial community. *The ISME journal*, **4**, 337-345, doi: 10.1038/ismej.2009.122.

Eastwell, K., Sholberg, P. & Sayler, R. 2006. Characterizing potential bacterial biocontrol agents for suppression of *Rhizobium vitis*, causal agent of crown gall disease in grapevines. *Crop Protection,* **25**, 1191 – 1200, doi: 10.1016/j.cropro.2006.03.004

Ferrenberg, S., O'Neill, S. P., Knelman, J. E., Todd, B., Duggan, S., Bradley, D., Robinson, T., Schmidt, S. K., Townsend, A. R., Williams, M. W., Cleveland, C. C., Melbourne, B. A., Jiang, L. & Nemergut, D. R. 2013. Changes in assembly processes in soil bacterial communities a wildfire disturbance. The *ISME Journal*, **7**, 1102–1111, doi: 10.1038/ismej.2013.11

Fierer, N. & Jackson, R. B. 2006. The diversity and biogeography of soil bacterial communities. *Proceedings of the National Academy of Sciences*, **103**, 626-631. doi: 10.1073/pnas.0507535103.

Fierer, N., Bradford, M. A. & Jackson, R. B. 2007. Toward an ecological classification of soil bacteria. *Ecology*, **88**, doi: 10.1890/05-1839.

Fierer, N. 2017. Embracing the unknown: disentangling the complexities of the soil microbiome. *Nature Reviews Microbiology*, **15**, 579–590, doi: 10.1038/nrmicro.2017.87.

Fierer, N., Bradford, M. A., Jackson, R. B. 2007. Toward an ecological classification of soil bacteria. *Ecology*, **88**, 1354-1364, doi: 10.1890/05-1839.

Fierer, N. & Jackson, R. B. 2006. The diversity and biogeography of soil bacterial communities. *Proceedings of National Academy of Sciences*, **103**, 626-631, doi: 10.1073/pnas.0507535103.

Fischer, D., Pfitzner, B., Schmid, M., Simões-Araújo, J. L., Reis, V. M., Pereira, W., Ormeño-Orrillo, E., Hai, B., Hofmann, A., Schloter, M., Martinez-Romero, E., Baldani, J. I. & Hartmann, A. 2012. Molecular characterisation of the diazotrophic bacterial community in uninoculated and inoculated field-grown sugarcane (Saccharum sp.). *Plant and Soil*, **356**, 83-99, doi: 10.1007/s11104-011-0812-0.

Frantzeskakis, L., Di Pietro, A., Rep, M., Schirawski, J., Wu, C.-H. & Panstruga, R. 2020. Rapid evolution in plant–microbe interactions – a molecular genomics perspective. *New Phytologist*, **225**, 1134-1142, doi:10.1111/nph.15966.

Freedman, Z. & Zak, D. R. 2015. Soil bacterial communities are shaped by temporal and environmental filtering: evidence from a long-term chronosequence. *Environmental Microbiology*, **17**, doi: 10.1111/1462-2920.12762.

Fry, E. L., De Long, J. R., Garrido, L. Á., Alvarez, N., Carrillo, Y., Castañeda-Gómez, L., Chomel, M., Dondini, M., Drake, J. E., Hasegawa, S., Hortal, S., Jackson, B. G., Jiang, M., Lavallee, J. M., Medlyn, B. E., Rhymes, J., Singh, B. K., Smith, P., Anderson, I. C., Bardgett, R. D., Baggs, E. M. & Johnson, D. 2019. Using plant, microbe, and soil fauna traits to improve the predictive power of biogeochemical models. *Methods in Ecology and Evolution*, **10**, 146-157, doi: 10.1111/2041-210X.13092.

Fuhrman, J. A. 2012. Metagenomics and its connection to microbial community organization. *F1000 Biology Reports*, **4**, doi: 10.3410/B4-15PMID:22912649.

Fuhrman, J. A., Cram, J. A., & Needham, D. M. 2015. Marine microbial community dynamics and their ecological interpretation. *Nature Reviews Microbiology*, **13**, 133-146, doi: 10.1038/nrmicro3417.

Gaiero, J. R., McCall, C. A., Thompson, K. A., Day, N. J., Best, A. S. & Dunfield, K. E. 2013. Inside the root microbiome: Bacterial root endophytes and plant growth promotion. *American Journal of Botany*, **100**, 1738-1750, doi: 10.3732/ajb.1200572.

Geisen, S., Wall, D. H. & van der Puttern, W. H. 2019. Challenges and Opportunities for Soil Biodiversity in the Anthropocene. *Current Biology*, **29**, doi: 10.1016/j.cub.2019.08.007.

Glick, B. R. 2005. Modulation of plant ethylene levels by the bacterial enzyme ACC deaminase. *FEMS Microbiology Letters*, **251**, 1-7, doi: 10.1016/j.femsle.2005.07.030.

Greenhalgh, K., Meyer, K. M., Aagaard, K. M. & Wilmes, P. 2016. The human gut microbiome in health: establishment and resilience of microbiota over a lifetime. *Environmental Microbiology*, **18**, doi: 10.1111/1462-2920.13318.

Griffiths, B. S. & Philippot, L. 2013. Insights into the resistance and resilience of the soil microbial community. *FEMS Microbiology Reviews*, **37**, 112-129, doi: 10.1111/j.1574-6976.2012.00343.x.

Halfvarson, J., Brislawn, C., Lamendella, R., Vázquez-Baeza, Y., Walters, W. A., Bramer, L. M., D'Amato, M., Bonfiglio, F., McDonald, D., Gonzalez, A., McClure, E. E., Dunklebarger, M. F., Knight, R., Jansson, J. K. 2017. Dynamics of the human gut microbiome in inflammatory bowel disease. *Nature Microbiology*, **2**, doi: 10.1038/nmicrobiol.2017.4

Hartmann, M., Niklaus P. A., Zimmermann, S., Schmutz, S., Kremer, J., Abarenkov, K., Lüscher, P., Widmer, F. & Frey, B. 2014. Resistance and resilience of the forest soil microbiome to logging-associated compaction. *The ISME Jornal,* **8,** 226–244, doi: 10.1038/ismej.2013.

Harun-Or-Rashid, Md. & Chung, Y. R. Induction of Systematic Resistance against Insect Herbivores in Plants by Beneficial Soil Microbes. 2017. *Frontiers in Plant Science*, **8**, doi: 10.3389/fpls.2017.01816.

Hassan, S. & Mathesius, U. 2012. The role of flavonoids in root-rhizosphere signalling: opportunities and challenges for improving plant-microbe interactions. *Journal of Experimental Botany*, **63**, 3429-3444, doi: 10.1093/jxb/err430.

Herzog, S., Wemheuer, F., Wemheuer, B. & Daniel, R. 2015. Effects of Fertilization and Sampling Time on Composition and Diversity of Entire and Active Bacterial Communities in German Grassland Soils. *PLoS ONE*, **10,** doi: 10.1371/journal.pone.0145575

Hiltner, L. 1904. Ueber neuere Erfahrungen und Probleme auf dem Gebiete der Bodenbakteriologie und unter besonderer BerUcksichtigung der Grundungung und Brache. *Arbeiten der Deutschen Landwirtschafts Gesellschaft*, **98**, 59-78.

Hodgson, D., McDonald, J. L. & Hosken, D. J. 2015. What do you mean, 'resilient'? *Trends in Ecology and Evolution*, **30**, 503-506, doi: 10.1016/j.tree.2015.06.010.

Hubbell, S. P. 2001. The Unified Neutral Theory of Biodiversity and Biogeography. Monographs in Population Biology, Vol 32. Princeton University Press: Princeton, USA.

Hug, L. A., Baker, **B. J., Anantharaman, K., Brown, C. T., Probst, A. J., Castelle, C. J., Butterfield, C. N., Hernsdorf, A. W., Amano, Y., Ise, K., Suzuki, Y., Dudek, N., Relman, D. A., Finstad, K. M., Amundson, R., Thomas, B. C. & Banfield**, J. F. 2016. A new view of the tree of life. *Nature Microbiology*, **1,** doi: 10.1038/nmicrobiol.2016.48

James, E.K. 2000. Nitrogen fixation in endophytic and associative symbiosis. *Field Crops Research*, **65**, 197-209, doi: 10.1016/S0378-4290(99)00087-8.

Jansson, J. K. & Hofmockel, K. S. 2020. Soil microbiomes and climate change. *Nature Reviews Microbiology*, **18**, 35–46, doi: 10.1038/s41579-019-0265-7.

Jiao, S., Chen, W., Wang, E., Wang, J., Liu, Z., Li, Y. & Wei, G. 2016. Microbial succession in response to pollutants in batch-enrichment culture. *Scientific Reports*, **6**, doi: 10.1038/srep21791.

Jiao, S., Du, N., Zai, X., Gao, X., Chen, W. & Wei, G. 2019. Temporal dynamics of soil bacterial communities and multifunctionality are more sensitive to introduced plants than to microbial additions in a multicontaminated soil. *Land Degradation & Development*, **30**, doi: 10.1002/ldr.3272.

Johnson, J. S., Spakowicz, D. J., Hong, B., Petersen, L. M., Demkowicz, P., Chen, L., Leopold, S. R., Hanson, B. M., Agresta, H. O., Gerstein, M., Sodergren, E. & Weinstock, G. M. 2019. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nature Communications*, **10**, doi: 10.1038/s41467-019-13036-1.

Kaeberlein, T., Lewis, K. & Epstein, S. S. 2002. Isolating "uncultivable" microorganisms in pure culture in a simulated natural environment. *Science*, **296**, 1127-1129, doi: 10.1126/science.1070633.

Kaiser, K., Wemheuer, B., Korolkow, V., Wemheuer, F., Nacke, H., Schöning, I., Schrumpf, M. & Daniel, R. 2016. Driving forces of soil bacterial community structure, diversity, and function in temperate grasslands and forests. Scientific Reports, **6**, doi: 10.1038/srep33696.

Kamaruzzaman, M. A., Abdullah, S. R. S., Hasan, H. A., Hassan, M., Othman, A. R. & Idris, M. 2020. Characterisation of Pb-resistant plant growth-promoting rhizobacteria (PGPR) from *Scirpus grossus*. *Biocatalysis and Agricultural Biotechnology*, **23**, doi: 10.1016/j.bcab.2019.101456.

Kell, D. B. & Young, M. 2000. Bacterial dormancy and culturability: the role of autocrine growth factors. *Current Opinion in Microbiology*, **3**, 238-243, doi: 10.1016/s1369-5274(00)00082-5.

Kitanov, G. M. 2001. Hypericin and pseudohypericin in some *Hypericum* species. *Biochemical Systematics and Ecology*, **29**, 171–178, doi: 10.1016/S0305-1978(00)00032-6.

Lladó, S., Žifčáková, L., Větrovský, T., Eichlerová, I. & Baldrian, P. 2016. Functional screening of abundant bacteria from acidic forest soil indicates the metabolic potential of Acidobacteria subdivision 1 for polysaccharide decomposition. *Biology and Fertility of Soils*, **52,** 251–260, doi: 10.1007/s00374-015-1072-6.

Kinrys, G., Coleman, E. & Rothstein, E. 2009. Natural remedies for anxiety disorders: potential use and clinical applications. *Depression and Anxiety*. **26**, 259-265, doi: 10.1002/da.20460.

Kirchmann, H. & Gerzabek, M. H. 1999. Relationship between soil organic matter and micropores in a long-term experiment at Ultuna, Sweden*. Journal of Plant Nutrition and Soil Science*, **162**, 493-498, doi: 10.1002/(SICI)1522-2624(199910)162:53.0.CO;2-S.

Kitanov, G. M. 2001. Hypericin and pseudohypericin in some *Hypericum* species. *Biochemical Systematics and Ecology*, **29**, 171–178, doi: 10.1016/S0305-1978(00)00032-6.

Klemow, K. M., Bartlow, A., Crawford, J., Kocher, N., Shah, J. & Ritsick M. 2011. Chapter 11 - Medical Attributes of St. John's Wort (Hypericum perforatum). In: Benzie, I. F. F. & Wachtel-Galor, S. (eds) *Herbal Medicine: Biomolecular and Clinical Aspects (2nd edition)*. CRC Press/Taylor & Francis, Boca Raton (FL), {ISBN} 978-1-4398-0713-2, doi: https://www.ncbi.nlm.nih.gov/books/NBK92750/.

Knelman, J. E., Legg, T. M., O'Neill, S. P., Washenberger, C. L., González, A., Cleveland, C. C. & Nemergut, D. R. 2012. Bacterial community structure and function change in association with colonizer plants during early primary succession in a glacier forefield. *Soil Biology and Biochemistry*, **46**, 172–180, doi: 10.1016/j.soilbio.2011.12.001.

Košuth, J., Smelcerovic, A., Borsch, T., Zuehlke, S., Karppinen, K., Spiteller, M., Hohtola, A. & Čellárová, E. 2011. The hyp-1 gene is not a limiting factor for hypericin biosynthesis in the genus Hypericum. *Functional Plant Biology*, **38**, 35-43, doi: 10.1071/fp10144.

Kusari, S. Zühlke, S., Borsch, T. & Spiteller, M. 2009. Positive correlations between hypericin and putative precursors detected in the quantitative secondary metabolite spectrum of Hypericum. *Phytochemistry*. **70**, 1222-1232. doi: 10.1016/j.phytochem.2009.07.022.

Kuypers, M., Marchant, H. & Kartal, B. 2018. The microbial nitrogen-cycling network. *Nature Reviews Microbiology*, **16**, 263–276, doi: 10.1038/nrmicro.2018.9.

Kuzyakov, Y. & Blagodatskaya, E. 2015. Microbial hotspots and hot moments in soil. Concept & review. *Soil Biology and Biochemistry*, **83**, 184-199, doi: 10.1016/j.soilbio.2015.01.025.

Landesman, W. J., Nelson, D. M. & Fitzpatrick, M. C. 2014. Soil properties and tree species drive ß-diversity of soil bacterial communities. *Soil Biology and Biochemistry*, **76**, 201–209, doi: 10.1016/j.soilbio.2014.05.025.

Lauber, C. L., Hamady, M., Knight, R. & Fierer, N. 2009. Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. *Applied and Environmental Microbiology*, **75,** 5111-5120, doi: 10.1128/AEM.00335-09.

Leibold, M. A. & McPeek, M. A. 2006. Coexistence of the niche and neutral perspectives in community ecology. *Ecology*, **87**, 1399–1410, doi: 10.1890/0012-9658(2006)87[1399:COTNAN]2.0.CO;2.

Lennon, J. & Jones, S. 2011. Microbial seed banks: the ecological and evolutionary implications of dormancy. *Nature Reviews Microbiology*, **9**, 119–130, doi: 10.1038/nrmicro2504.

Li, H., Su, J.-Q, Yang, X.-R. & Zhu, Y.-G. 2019. Distinct rhizosphere effect on active and total bacterial communities in paddy soils. *Science of The Total Environment*, **649**, 422-430, doi: 10.1016/j.scitotenv.2018.08.373.

Li, S., Tan, J., Yang, X., Ma, C. & Jiang, L. 2019. Niche and fitness differences determine invasion success and impact in laboratory bacterial communities. *The ISME Journal*, **13**, 402–412, doi: 10.1038/s41396-018-0283-x.

Liu, H., Carvalhais, L. C., Crawford, M., Singh, E., Dennis, P. G., Pieterse, C. M. J. & Schenk, P. M. 2017. Inner Plant Values: Diversity, Colonization and Benefits from Endophytic Bacteria. *Frontiers in Microbiology*, **8**, doi: 10.3389/fmicb.2017.02552

Liu, Y., Luo, M., Ye, R., Huang, J., Xiao, L., Hu, Q., Zhu, A. & Tong, C. 2019. Impacts of the rhizosphere effect and plant species on organic carbon mineralization rates and pathways, and bacterial community composition in a tidal marsh. *FEMS Microbiology Ecology*, **95**, doi: https://doi.org/10.1093/femsec/fiz120.

Ling, N., Zhang, W., Wang, D., Mao, J., Huang, Q., Guo, S. & Shen, Q. 2013. Root exudates from grafted-root watermelon showed a certain contribution in inhibiting *Fusarium oxysporum* f. sp. *niveum*. *PLoS ONE*, **8**, doi: 10.1371/journal.pone.0063383.

Lladó, S., López-Mondéjar, R. & Baldrian, P. 2017. Forest soil bacteria: diversity, involvement in ecosystem processes, and response to global change. *Microbiology and Molecular Biology Reviews*, **81**, e00063-16, doi: 10.1128/MMBR.00063-16.

Loeppmann, S., Semenov, M., Kuzyakov, Y. & Blagodatskaya, E. 2018. Shift from dormancy to microbial growth revealed by RNA:DNA ratio. *Ecological Indicators*, **85**, 603-612, doi: 10.1016/j.ecolind.2017.11.020.

Lopes, L. D., Pereira e Silva, M. de C. & Andreote, F. D. 2016. Bacterial abilities and adaptation toward the rhizosphere colonization. *Frontiers in Microbiology*, **7**, doi: 10.3389/fmicb.2016.01341.

Lundberg, D. S., Lebeis, S. L., Paredes, S. H., Yourstone, S., Gehring, J., Malfatti, S., Tremblay, J., Engelbrektson, A., Kunin, V., del Rio, T. G., Edgar, R. C., Eickhorst, T., Ley, R. E., Hugenholtz, P., Tringe, S. G. & Dangl, J. L. 2012. Defining the core Arabidopsis thaliana root microbiome. *Nature*, **488**, 86-90, doi: 10.1038/nature11237.

Mañero, F. J. G., Algar, E., Gómez, M. S. M., Sierra, M. D. S. & Solano, B. R. 2012. Elicitation of secondary metabolism in Hypericum perforatum by rhizosphere bacteria and derived elicitors in seedlings and shoot cultures. *Pharmaceutical Biology*, **50**, 1201-1209, doi: 10.3109/13880209.2012.664150.

Maron, P-A., Sarr, A., Kaisermann, A., Lévêque, J., Mathieu, O., Guigue, J., Karimi, B., Bernard, L., Dequiedt, S., Terrat, S., Chabbi, A. & Ranjard, L. 2018. *Applied and Environmental Microbiology*, **84**, e02738-e02717, doi:10.1128/AEM.02738-17.

Maróti, G. & Kondorosi, É. 2014. Nitrogen-fixing *Rhizobium*-legume symbiosis: are polyploidy and host peptide-governed symbiont differentiation general principles of endosymbiosis? *Frontiers in Microbiology*, **5**, doi: 10.3389/fmicb.2014.00326.

McCarthy A. 2010. Third generation DNA sequencing: pacific biosciences' single molecule real time technology. *Chemistry & Biology*, **17**, 675-676, doi: 10.1016/j.chembiol.2010.07.004.

Meyer, KM., Petersen, IAB., Tobi, E., Korte, L. & Bohannan, BJM. 2019. Use of RNA and DNA to Identify Mechanisms of Bacterial Community Homogenization. *Frontiers in Microbiology,* **10**, doi: 10.3389/fmicb.2019.02066

Mitchell, R. J., Campbell, C. D., Chapman, S. J. & Cameron, C. M. 2010. The ecological engineering impact of a single tree species on the soil microbial community. *Journal of Ecology*, **98**, doi: 10.1111/j.1365-2745.2009.01601.x.

Monteiro, R. A., Balsanelli, E., Wassem, R., Marin, A. M., Brusamarello-Santos, L. C. C., Schmidt, M. A., Tadra-Sfeir, M. Z., Pankievicz, V. C. S., Cruz, L. M., Chubatsu, L. S., Pedrosa, F. O. & Souza, E. M. Herbaspirillum-plant interactions: microscopical, histological and molecular aspects. *Plant and Soil*, **356**, 175-196, doi: 10.1007/s11104-012-1125-7.

Morris, S. J. & Blackwood, C. B. 2015. The ecology of the soil biota and their function. *In: Soil Microbiology, Ecology and Biochemistry,* 273-309, doi: 10.1016/B978-0-12-415955-6.00010-4.

Nguyen, C. 2003. Rhizodeposition of organic C by plants: mechanisms and controls. *Agronomie*, **23**, 375-396, doi: 10.1051/agro:2003011.

Nielsen, U. N., Ayres, E., Wall, D. H. & Bardgett, R. D. 2011. Soil biodiversity and carbon cycling: A review and synthesis of studies examining diversity-function relationships. *European Journal of Soil Science*, **62**, 105– 116, doi: 10.1111/j.1365-2389.2010.01314.x.

Nürk, N. M., Madriñán, S., Carine, M. A., Chase, M. W. & Blattner, F. R. 2012. Molecular phylogenetics and morphological evolution of St. John's wort (*Hypericum*; Hypericaceae). *Molecular Phylogenetics and Evolution*, **66**, 1-16, doi: 10.1016/j.ympev.2012.08.022.

Oades, J. M. 1993. The role of biology in the formation, stabilization and degradation of soil structure. *Geoderma*, **56**, 377-400, doi: 10.1016/0016-7061(93)90123-3.

Oh, Y. M., Kim, M., Lee-Cruz, L., Lai-Hoe, A., Go, R., Ainuddin, N., Rahim, R. A., Shukor, N. & Adams, J. M. 2012. Distinctive bacterial communities in the rhizoplane of four tropical tree species. *Microbial Ecology*, **64**, 1018-1027, doi: 10.1007/s00248-012-0082-2.

Olanrewaju, O. S., Ayangbenro, A. S., Glick, B. R. & Babalola, O. O. 2019. Plant health: feedback effect of root exudates-rhizobiome interactions. *Applied Microbiology and Biotechnology*, **103**, 1155-1166, doi: 10.1007/s00253-018-9556-6.

Orgiazzi, A., Bardgett, R., Barrios, E., Behan-Pelletier, V., Briones, M. J. I., Chotte, J.-L., de Deyn, G. B., Eggleton, P., Fierer, N., Fraser, T., Hedlund, K., Jeffery, S., Johnson, N. C., Jones, A., Kandeler, E., Kaneko, N., Lavelle, P., Lemanceau, P. Miko, L., Montanarella, L., Moreira, F. M. S., Ramirez, K. S., Scheu, S., Singh, B. K., Six, J., van der Putten, W. H. & Wall, D. H. (eds). 2016. Global Soil Biodiversity Atlas. *In: European Commission.* Publications Office of the, European Union, Luxembourg, pp. 176. doi*:* 10.2788/2613.

Osler, G. H. R. & Sommerkorn, M. 2007. Toward a complete soil C and N cycle: Incorporating the soil fauna. *Ecology*, **88**, 1611-1621, doi:10.1890/06-1357.1.

Overmann, J. 2013. Chapter 7. Principles of enrichment, isolation, cultivation, and preservation of prokaryotes. *In: The Prokaryotes: Prokaryotic Biology and Symbiotic Associations*. 149–207, doi: 10.1007/978-3642-30194-0_7.

Overmann, J., Abt, B. & Sikorski, J. 2017. Present and future of culturing bacteria. *Annual Review of Microbiology,* **71**, 711-730, doi: 10.1146/annurev-micro-090816-093449.

Pacifico, D., Squartini, A., Crucitti, D., Barizza, E., Lo Schiavo, F., Muresu, R., Carimi, F. & Zottini, M. 2019. *Frontiers in Plant Science*, doi: 10.3389/fpls.2019.01256.

Peiffer, J. A., Spor, A., Koren, O., Jin, Z., Tringe, S. G., Dangl, J. L., Buckler, E. S. & Ley, R. E. 2013. Diversity and heritability of the maize rhizosphere microbiome under field conditions. *Proceedings of the National Academy of Sciences*, **110**, 6548-6553, doi: 10.1073/pnas.1302837110.

Pham, V. H. T. & Kim, J. 2012. Cultivation of unculturable soil bacteria. *Trends in Biotechnology*, 30, 475-484, doi: 10.1016/j.tibtech.2012.05.007.

Podolich, O., Ardanov, P., Zaets, I., Pirttilä, A. M. & Kozyrovska, N. 2015. Reviving of the endophytic bacterial community as a putative mechanism of plant resistance. *Plant and Soil,* **388**, doi; 10.1007/s11104-014-2235-1.

Puspita, I. D., Kamagata, Y., Tanaka, M., Asano, K. & Nakatsu, C. H. 2012. Are Uncultivated Bacteria Really Uncultivable? *Microbes and Environments*, **27**, 356-366, doi: 10.1264/jsme2.ME12092.

Quince, C., Walker, A. W., Simpson, J. T., Loman, N. J. & Segata, N. 2017. Shotgun metagenomics, from sampling to analysis. *Nature Biotechnology*, **35**, 833-844, doi: 10.1038/nbt.3935.

Raclariu, A. C., Paltinean, R., Vlase, L., Labarre, A., Manzanilla, V., Ichim, M. C., Crisan, G., Brysting, A. K. & de Boer, H. 2017. Comparative authentication of *Hypericum perforatum* herbal products

using DNA metabarcoding, TLC and HPLC-MS. *Scientific Reports*, **7**, doi: 10.1038/s41598-017-01389-w.

Rappé, M. S., Connon, S. A., Vergin, K. L. & Giovannoni, S. J. 2002. Cultivation of the ubiquitous SAR11 marine bacterioplankton clade. *Nature*, **418**, 630-633, doi: 10.1038/nature00917.

Raynaud, X. & Nunan, N. 2014. Spatial ecology of Bacteria at the Microscale in Soil. *PLoS ONE*, **9**, e87217, doi: 10.1371/journal.pone.0087217.

Reich, P. B., Oleksyn, J., Modrzynski, J., Mrozinski, P., Hobbie, S. E., Eissenstat, D. M., Chorover, J., Chadwick, O. A., Hale, C. M. & Tjoelker, M. G. 2005. Linking litter calcium, earthworms and soil properties: a common garden test with 14 tree species. *Ecology Letters*, **8**, doi: 10.1111/j.1461-0248.2005.00779.x.

Rillig, M. C., Ryo, M., Lehmann, A., Aguilar-Trigueros, C. A., Buchert, S. Wulf, A., Iwasaki, A., Roy, J. & Yang, G. 2019. The role of multiple global change factors in driving soil functions and microbial biodiversity. *Science*, **366**, doi: 10.1126/science.aay2832.

Roesch, L. F. W., Fulthorpe, R. R., Riva, A., Casella, G., Hadwin, A. K. M., Kent, A. D., Daroub, S. H., Camargo, F. A. O., Farmerie, W. G. & Triplett, E. W. 2007. Pyrosequencing enumerates and contrasts soil microbial diversity. *The ISME Journal*, **1**, 283–290, doi: 10.1038/ismej.2007.53.

Romanowicz, K. J., Freedman, Z. B., Upchurch, R. A., Argiroff, W. A & Zak, D. R. 2016. Active microorganisms in forest soils differ from the total community yet are shaped by the same environmental factors: the influence of pH and soil moisture. *FEMS Microbiology Ecology*, **92**, doi: 10.1093/femsec/fiw149.

Ryu, C. M. 2009. Rhizosphere bacteria help plants tolerate abiotic stress. *Trends in Plant Science,* **14**, doi: 0.1016/j.tplants.2008.10.004.

Sait, M., Hugenholtz, P. & Janssen, P. H. 2002. Cultivation of globally distributed soil bacteria from phylogenetic lineages previously only detected in cultivation-independent surveys. *Environmental Microbiology*, **4**, 654-666, doi: 10.1046/j.1462-2920.2002.00352.x.

Sessitsch, A., Hardoim, P., Doring, J., Weilharter, A., Krause, A., Woyke, T., Mitter, B., Hauberg-Lotte, L., Friedrich, F., Rahalkar, M., Hurek, T., Sarkar, A., Bodrossy, L., van Overbeek, L., Brar, D., van Elsas, J. D. & Reinhold-Hurek, B. 2012. Functional characteristics of an endophyte community colonizing rice roots as revealed by metagenomic analysis. *Molecular Plant-Microbe Interactions*, **25**, 28-36, doi: 10.1094/MPMI-08-11-0204.

Shen, C, Ni, Y., Liang, W., Wang, J. & Chu, H. 2015. Distinct soil bacterial communities along a small-scale elevational gradient in alpine tundra. *Frontiers in Microbiology*, **6**, doi: 10.3389/fmicb.2015.00582.

Singh, D., Lee-Cruz, L., Kim, W. S., Kerfahi, D., Chun, J. H. & Adams, J. M. 2014. Strong elevational trends in soil bacterial community composition on Mt. Halla, South Korea. *Soil Biology and Biochemistry*, **68**, 140–149, doi: 10.1016/j.soilbio.2013.09.027.

Singh, B. & Satyanarayana, T. 2011. Microbial phytases in phosphorus acquisition and plant growth promotion. *Physiology and Molecular Bioloy of Plants*, **17**, 93 – 103, doi: 10.1007/s12298-011-0062-x.

Smalla, K., Wieland, G., Buchner, A., Zock, A., Parzy, J., Kaiser, S., Roskot, N., Heuer, H. & Berg, G. 2001. Bulk and rhizosphere soil bacterial communities studied by denaturing gradient gel

electrophoresis: plant-dependent enrichment and seasonal shifts revealed. *Applied and Environmental Microbiology*, **67**, 4742-4751, doi: 10.1128/AEM.67.10.4742-4751.2001.

Smelcerovic, A. & Spiteller, M. 2006. Phytochemical analysis of nine Hypericum L. species from Serbia and the F.Y.R. Macedonia. *Pharmazie*, **61**, 251-252.

Staley, J. T. & Konopka, A. 1985. Measurement of in situ activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. *Annual Review of Microbiology*, **39**, 321-346, doi: 10.1146/annurev.mi.39.100185.001541.

Stewart, E. J. 2012. Growing Unculturable Bacteria. *Journal of Bacteriology*, **194**, 4151-4160, doi: 10.1128/JB.00345-12.

Stojanović, G., Đorđević, A. & Šmelcerović, A. 2013. Do other Hypericum species have medical potential as St. John's wort (Hypericum perforatum)? *Current Medicinal Chemistry*, **20**, 2273 – 2295, doi: 10.2174/0929867311320180001

Suarez-Moreno, Z. R., Caballero-Mellado, J., Coutinho, B. G., Mendonca-Previato, L., James, E. K. & Venturi, V. 2012. Common features of environmental and potentially beneficial plant-associated Burkholderia. *Microbial Ecology*, **63**, 249-266, doi: 10.1007/s00248-011-9929-1.

Sugiyama, A. 2019. The soybean rhizosphere: Metabolites, microbes, and beyond – A review. *Journal of Advanced Research*, **19**, doi:10.1016/j.jare.2019.03.005.

Tanaka, Y., Hanada, S., Manome, A., Tsuchida, T., Kurane, R., Nakamura, K. & Kamagata, Y. 2004. *Catellibacterium nectariphilum* gen. nov., sp. nov., which requires a diffusible compound from a strain related to the genus Sphingomonas for vigorous growth. *International Journal of Systematic and Evolutionary Microbiology*, **54**, 955-959, doi:10.1099/ijs.0.02750-0.

Tardy, V., Spor, A., Mathieu, O., Lévèque, J., Terrat, S., Plassart, P., Regnier, T., Bardgett, R. D., van der Putten, W. H., Roggero, P. P., Seddaiu, G., Bagella, S., Lemanceau, P., Ranjard, L. & Maron, P-A. 2015. Shifts in microbial diversity through land use intensity as drivers of carbon mineralization in soil. *Soil Biology and Biochemistry*, **90,** 204-213, doi: 10.1016/j.soilbio.2015.08.010.

Thompson, L. R., Sanders, J. G., McDonald, D., Amir, A., Ladau, J., Locey, K. J., Prill, R. J., Tripathi, A., Gibbons, S. M., Ackermann, G., Navas-Molina, J. A., Janssen, S., Kopylova, E., Vázquez-Baeza, Y., González, A., Morton, J. T., Mirarab, S., Xu, Z. Z., Jiang, L., Haroon, M. F., Kanbar, J., Zhu, Q., Song, S. J., Kosciolek, T., Bokulich, N. A., Lefler, J., Brislawn, C. J., Humphrey, G., Owens, S. M., Hampton-Marcell, J., Berg-Lyons, D., McKenzie, V., Fierer, N., Fuhrman, J. A., Clauset, A., Stevens, R. L., Shade, A., Pollard, K. S., Goodwin, K. D., Jansson, J. K., Gilbert, J. A., Knight, R. & & The Earth Microbiome Project Consortium. 2017. A communal catalogue reveals Earth's multiscale microbial diversity. *Nature*, **551**, 457-463, doi: 10.1038/nature24621.

Tringe, S. G. & Hugenholtz, P. 2008. A renaissance for the pioneering 16S rRNA gene. *Current Opinion in Microbiology*, **11**, 442-446, doi: 10.1016/j.mib.2008.09.011.

Turner, T. R., James, E. K. & Poole, P. S. 2013. The plant microbiome. *Genome Biology*, **14**, doi: 10.1186/gb-2013-14-6-209.

Vacheron, J., Desbrosses, G., Bouffaud, M.-L., Touraine, B., Moënne-Loccoz, Y., Muller, D., Legendre, L., Wisniewski-Dyé, F. & Prigent-Combaret, C. 2013. Plant growth-promoting rhizobacteria and root system functioning. *Frontiers in Plant Science*, **4**, doi: 10.3389/fpls.2013.00356.

van Elsas, J. D. 2019. Chapter 1 - The soil environment. *In*: van Elsas, J. D., Trevors, J. T., Soares Rosado, A. & Nannipieri, P. (eds) *Modern Soil Microbiology, Third Edition*. CRC Press, Boca Raton, 3-18, doi: 10.1201/9780429059186.

Vartoukian, S. R., Palmer, R. M. & Wade, W. G. 2010. Strategies for culture of 'unculturable' bacteria. *FEMS Microbiology Letters*, **309**, 1-7, doi: 10.1111/j.1574-6968.2010.02000.x**.**

Vieira, S., Sikorski, J., Dietz, S. Herz, K., Schrumpf, M., Bruelheide, H., Scheel, D., Friedrich, M. W. & Overmann, J. 2020. Drivers of the composition of active rhizosphere bacterial communities in temperate grasslands. *The ISME Journal*, **14,** 463-475, doi: 10.1038/s41396-019-0543-4.

Vorholt, J. A. 2012. Microbial life in the phyllosphere. *Nature Reviews Microbiology*, **10**, 828-840, doi: 10.1038/nrmicro2910.

Walker, T. S., Bails, H. P. & Vivanco, J. M. 2002. Jasmonic acid-induced hypericin production in cell suspension cultures of *Hypericum perforatum* L. (St. John's wort). *Phytochemistry*, **60**, 289-293, doi: 10.1016/s0031-9422(02)00074-2.

Wall, D. H., Nielsen, U. N. & Six, J. 2015. Soil biodiversity and human health. *Nature*, **528**, 69-76, doi: 10.1038/nature15744.

Wang, G., Li, L., Wang, X., Li, X., Zhang, Y., Yu, J., Jiang, J., You, X. & Xiong, Y. Q. 2019. Hypericin enhances ß-lactam antibiotics activity by inhibiting sarA expression in methicillin-resistant *Staphylococcus aureus. Acta Pharmaceutica Sinica B,* **9**, doi: 10.1016/j.apsb.2019.05.002.

Wagg, C., Schlaeppi, K., Banerjee, S., Kuramae, E. E. & van der Heijden, M. G. A. 2019. Fungal-bacterial diversity and microbiome complexity predict ecosystem functioning. *Nature Communications,* **10,** doi: 10.1038/s41467-019-12798-y.

Wieland, G., Neumann, R. & Backhaus, H. 2001. Variation of microbial communities in soil, rhizosphere, and rhizoplane in response to crop species, soil type, and crop development. *Applied and Environmental Microbiology,* **67**, 5849-5854, doi: 10.1128/AEM.67.12.5849-5854.2001.

Wu, L., McCluskey, K., Desmeth, P., Liu, S., Hideaki, S., Yin, Y., Moriya, O., Itoh, T., Kim, C.Y., Lee, J.-S., Zhou, Y., Kawasaki, H., Hazbón, M. H., Robert, V., Boekhout, T., Lima, N., Evtushenko, L., Boundy-Mills, K., Bunk, B., Moore, E. R. B., Eurwilaichitr, L., Ingsriswang, S., Shah, H., Yao, S., Jin, T., Huang, J., Shi, W., Sun, Q., Fan, G., Li, W., Li, X., Kurtböke, İ. & Ma, J. 2018. The global catalogue of microorganisms 10K type strain sequencing project: closing the genomic gaps for the validly published prokaryotic and fungi species. *GigaScience*, **7**, doi: 10.1093/gigascience/giy026.

Yang, J., Kloepper, J. & Ryu, C. M. 2009. Rhizosphere bacteria help plants tolerate abiotic stress. *Trends in Plant Science*, **14**, doi; 10.1016/j.tplants.2008.10.004.

Yuan, Y. L., Si, G. C., Wang, J., Luo, T. X. & Zhang, G. X. 2014. Bacterial community in alpine grasslands along an altitudinal gradient on the Tibetan Plateau. *FEMS Microbiology Ecology*, **87**, 121-132, doi: 10.1111/1574-6941.12197.

Zarraonaindia, I., Smith, D. P. & Gilbert, J. A. 2013. Beyond the genome: community-level analysis of the microbial world. *Biology & Philosophy*, **28,** 261-282, doi: 10.1007/s10539-012-9357-8.

Zhalnina, K., Dias, R., de Quadros, P. D., Davis-Richardson, A., Camargo, F. A. O., Clark, I. M., McGrath, S. P., Hirsch, P. R. & Triplett, E. W. 2015. Soil pH determines microbial diversity and composition in the park grass experiment, *Microbial Ecology*, **69**, 395-406, doi: 10.1007/s00248-014-0530-2.

Zhang, X. F., Xu, S. J., Li, C. M., Zhao, L., Feng, H. Y., Yue, G. Y., Ren, Z. W. & Cheng, G. D. 2014.The soil carbon/nitrogen ratio and moisture affect microbial community structures in alkaline permafrost-affected soils with different vegetation types on the Tibetan Plateau. *Research in Microbiology*, **165**, 128-139, doi:10.1016/j.resmic.2014.01.002

Zhou, J., He, Z., Yang, Y., Deng, Y., Tringe, S. G. & Alvarez-Cohen, L. 2015. High-Throughput Metagenomic Technologies for Complex Microbial Community Analysis: Open and Closed Formats. *MBio*, **6**, e02288-14, doi: 10.1128/mBio.02288-14.

Chapter 3 – Material and Methods

Both culture-dependent and culture-independent techniques were employed to study bacterial communities of *Hypericum* plants, with a special focus on the infamous St. John's wort (*H. perforatum*) as the first step in understanding the connection between associated bacterial community and *Hypericum* unique secondary metabolites, namely hypericin and hyperforin.

3.1 – Bacterial community analysis of *Hypericum* species

The study involved two scoping studies, where the wild and already-cultivated *Hypericum* plants were taken from the fields and greenhouses of the Leibniz Institute of Plant Biochemistry (IPB) Halle in 2016 (first scoping study) and the Botanic Garden and Botanical Museum Berlin – Dahlem (BGBM) in 2018 (second scoping study), and a full-factorial crossed design study where the plants were grown from the seedlings under controlled conditions in one of the BGBM greenhouses in 2019. The scoping studies (both were merged and analysed together) were intended to investigate bacterial communities of *Hypericum* plants in their natural habitat and to identify the potential key players in the production of hypericin and hyperforin by comparing bacterial communities of hypericin and hyperforin-producing species and non-producer. However, the task was proven to be a great challenge due to the differences that were observed even in the bulk soil bacterial communities across samples from distinct sampling locations. Thus, the plant species effect was confounded with soil effect and could not be disentangled from a soil effect. This only allowed comparison between bulk soil and rhizosphere community of the plant with the same sampling origin. Therefore, a controlled greenhouse experiment with full-factorial crossed design was conducted to investigate the impact of habitat type (bulk soil, the rhizosphere, or inside the roots), soil substrate (with distinct pH), and plant species on *Hypericum* bacterial communities. Investigation of the total (DNA-based) and active (RNA-based) bacterial communities was performed and the information was subsequently used to identify active taxa based on rRNA:rDNA ratio that potentially trigger hypericin and hyperforin.

3.1.1 – Sampling for the scoping study

In 2016, three biological replicates of wild *H. perforatum* and *H. polyphyllum* (from different fields) and greenhouse-cultivated *H. perforatum* plants were sampled (the first scoping study). The rhizosphere was taken manually by hand and put into a 50 mL falcon tube for all biological replicates and defined as the soil that was still attached to the roots when the plants were taken out from the soil. On the other hand, bulk soil was partially sampled (only two biological replicates of wild *H. perforatum* and one biological replicate of *H. polyphyllum)*. Bulk soil of the greenhouse samples was

not available as the soil was strongly attached to the root, and thus is still a part of the rhizosphere, when the plant was taken out of the pot. The pH measurement was conducted in water and calcium chloride (2 mM $CaCl_2$).

Another sampling campaign took place in 2018 at the BGBM Berlin and referred as the second scoping study. In total, 6 biological replicates of each wild *H. perforatum* and *H. androsaemum* plants were sampled from two different fields of the BGBM. In addition, 5 biological replicates of the cultivated *H. balearicum* plants were sampled from one of the greenhouses. Bulk soil and the rhizosphere were taken from all biological replicates. Instead of collecting all soil that was attached to the roots after plants were taken out from the ground during rhizosphere collection (as had been done previously for samples collected for the first scoping study), manual shaking of the roots was perfomed subsequently to remove any excess soil. In addition, while samples of the first scoping study (retrieved from Halle) comprising only the rhizosphere, the roots were taken along with the rhizosphere and put into the same falcon tube for samples of the second scoping study (collected from Berlin). The rhizosphere and roots of samples belonging to the second scoping study were further mixed together in the laboratory by crushing both with the help of liquid nitrogen. Thus, the rhizosphere samples of the second scoping study also contained root (endophytic) community. The samples were kept in -80 ^{O}C in soil preservation solution (LifeGuard™) to conserve both DNA and RNA.

3.1.2 – A full-factorial crossed design, controlled-greenhouse experiment

3.1.2.1 – Experimental design

An experiment with a full-factorial crossed design was conducted in one of the BGBM greenhouses in 2019 to determine the important drivers of *Hypericum* bacterial communities (Figure 8). The plants were germinated for 2 weeks, grown in a standard soil (equal shares of garden soil and alkaline sand) for 7 weeks, and transferred to an alkaline (Vulkatree Humin; pH 7.5; granularity of 0-16 mm), neutral (Vulkatree V/P; pH 7; granularity of 0-12 mm) and acidic soils (Vulkaplus Intensiv; pH 5.5 to 5.6; granularity of 0-23 mm) (VulkaTec Riebensahm, GmbH, Kretz, Germany; https://www.vulkatec.de) and incubated for 14 weeks prior to harvest.

Cultivated *Hypericum* species included *H. perforatum, H. olympicum, H. balearicum, H. bithnycum, H. tetrapterum, H. delphicum, H. hirsutum, H. maculatum,* and *H. canariense*. The first three species were selected for the first investigation of bacterial communities of *Hypericum. H. perforatum* was selected as the plant was the most notable species for the production of hypericin and hyperforin. *H. olympicum* was selected as an additional hypericin and hyperforin-producing species (the plant

species produces lower amounts of the compounds compared to *H. perforatum,* albeit morphologically similar) whereas *H. balearicum* was chosen since it is a non-producer. Thus, comparison between producer and non-producer plants could be performed to find potential key players in the production of hypericin and hyperforin.

The alkaline and acidic soils were selected for the analysis, since both gave the most contrasting yield of plant biomass. Although approximately 20 biological replicates were available for each plant species, only 6 to 10 replicates were harvested for the first analysis. The rest would be harvested in the near future and could involve analysis of temporal dynamics of *Hypericum* bacterial communities. Altogether, 5 biological replicates for each plant species were investigated by culture-independent methods as it offers a good compromise between time and workload effort and statistical power for data analysis.

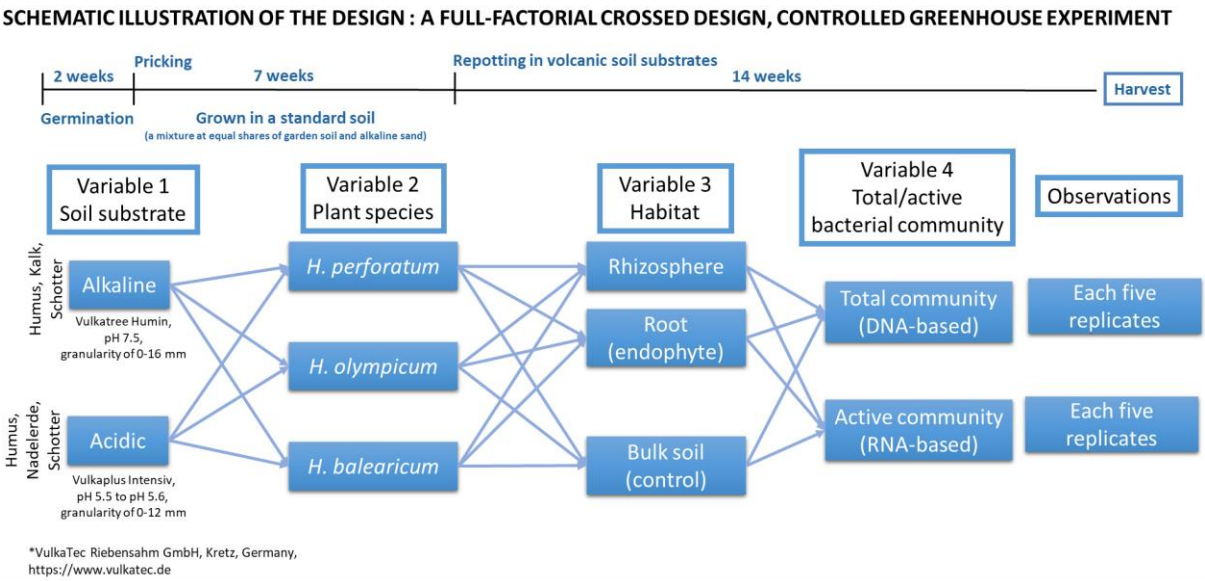


Figure 8 – The illustration of the full-factorial crossed design greenhouse experiment where the impact of soil substrate (acidic and alkaline), plant species (*H. perforatum, H. olympicum,* and *H. balearicum*) and habitat type (the rhizosphere, roots, and bulk soil) was determined. The plants were germinated for 2 weeks, transported to a standard soil and incubated for 7 weeks, and repotted in the volcanic soil substrates (acidic and alkaline) for 14 weeks prior to the harvest. Both total (DNA-based) and active (RNA-based) bacterial communities were investigated from 5 biological replicates of each plant species.

3.1.2.2 – The rhizosphere, roots, and bulk soil samples collection

Each plant was taken out of the pot and the soil that was not attached to the root after manual-shaking was collected as bulk soil. The rhizosphere was separated from the roots through a washing process with 0.9% NaCl solution containing 0.01% Tween 80 (Barillot *et al.* 2013) and thus rhizosphere community was assumed to contain rhizoplane community, albeit not further checked. The suspensions were centrifuged to collect the soil and the associated microbial cells. All samples were stored in liquid nitrogen before further processing. In total, 90 different samples were available including 3 different plant species (*H. perforatum, H. olympicum,* and *H. balearicum*), 2 distinct soil substrates (acidic and alkaline), and 3 habitat types (bulk, rhizosphere + rhizoplane, and root), each with 5 biological replicates.

3.1.3 – Co-extraction of RNA and DNA

Co-extraction of RNA and DNA was performed with the method described previously (Lueders *et al.* 2003) with some modifications according to Wüst *et al.* 2016. Each individual sample was thawed on ice and put into a 2 mL screw cap tube containing 0.7 g sterilized zirconium/silica beads with a diameter of 0.1 mm. A 750 µL of 120 nM Na_2HPO_4 buffer (pH 8) and 250 µL TNS buffer (Tris-HCl, sodium chloride, SDS) were added to the tube and beat beating was performed for 45 s at 6.5 $m \cdot s^{-1}$, followed by centrifugation for 20 min. Unless mentioned otherwise, centrifugation was done at maximum speed and at 4 °C. The supernatant was transferred and extracted with 1 volume (500 µL) of phenol/chloroform/isoamylalcohol (25:24:1) and 1 volume of chloroform/isoamylalcohol (24:1) with each step followed by a centrifugation for 5 min. The nucleic acid was further pelleted by mixing the supernatant with 2 volumes of PEG solution, followed by a centrifugation for 90 min. The pellets were washed with 500 µL ice-cold 70% EtOH and centrifuged for 30 min. The resulted co-extract was checked with the NanoDrop ND-100 (Peqlab Biotechnologie, Erlangen, Germany) and stored in 100 µL EB buffer at -80 °C.

3.1.4 – DNA digestion and reverse transcription to construct cDNA library

A 50 µL of the coextract was digested with RNAse free DNase I (Thermoscientific, Waltham, MA, U.S.) following the instructions of the manufacturer. The pellets were washed with 500 µL of ice-cold 70% EtOH, resuspended in 50 µL RNase free water and stored in -80 ^{O}C before further processing. The cDNA synthesis was performed using the GoScript Reverse Transcription System (Promega, Madison, WA, USA), according to the manufacturer instructions and with the employment of random hexamers.

3.1.5 – Indexing PCR by Bartram method

3.1.5.1 – cDNA Indexing

For cDNA library, the 16S rRNA V3 region was amplified in triplicates using modified primers 341F (5'- CCTACGGGWGGCWGCAG-3') and 515R (5'-CCGCGGCTGCTGGCAC-3') (Muyzer *et al.* 1993). The modified primer pairs contain an adapter sequence and the binding site of sequencing primer. In addition, the reverse primer also contains 6 Indexing-nucleotides (Bartram *et al.* 2011). Amplicon sequencing was conducted in a Veriti 96-well thermal cycler (Applied Biosystems, Foster city, CA, USA) in a total volume of 50 µL, which consisted of 10 µl PCR buffer (5x; GC Phusion buffer), 1 µl dNTPs (10 mM), 1.5 µl DMSO (100% v/v), 0.2 µl of each of forward and reverse primers (each 50 pmol/µL), 0.5 µl Phusion High Fidelity DNA Polymerase (2 U μl^{-1}; Thermo Scientific, Waltham, USA), and 20 µl cDNA library product. The PCR program was started with an initial denaturation step at 94 °C for 5 min, proceeded with 20 cycles of denaturation at 94 °C for 15 s, annealing at 59 °C for 15 s, and elongation at 72 °C for 15 s, and finally ended with a final elongation step at 72 °C for 7 min.

3.1.5.2 – DNA Indexing

The remainder 50 µl of the DNA/RNA coextract was used for the construction of the V3 region of 16S rRNA DNA library. Prior to Bartram PCR, pre-amplification targeting the V3 region was performed using primer pairs 341F (5'- CCTACGGGWGGCWGCAG-3') and 515R (5'-CCGCGGCTGCTGGCAC-3') (Muyzer *et al.* 1993). The PCR-reaction mix consisted of 10 µl PCR buffer (5x; GC Phusion buffer), 1 µl dNTP mix (each 10 mM), 3.5 µl $MgCl_2$ (50 mM), 1.5 µl DMSO (100% v/v), 1 µl BSA (20 mg/mL), 0.2 µl of each forward and reverse primers (each 50 pmol/µL), 2 µL Phusion Hot Start II DNA Polymerase (2 U µl-1; Thermo Scientific, Waltham, USA), and 1-5 µL of template DNA, and further added with PCR-H_2O to a final volume of 50 µl. An initial denaturation step at 98 °C for 30s, followed by 15 cycles of denaturation at 98 °C for 10 s, annealing at 59 °C for 30 s, and elongation at 72 °C for 30 S, and finalize with a final elongation at 72 °C for 7 min was set for the PCR. This pre-amplification PCR was proceeded with Bartram PCR, using modified primer pairs as mentioned in the cDNA indexing. The reaction mix was the same with the pre-amplification except the volume of DMSO (100% v/v), Phusion Hot Start II DNA Polymerase (2 U μl^{-1}; Thermo Scientific, Waltham, USA), and template DNA (the product of preamplification), which were 2.5 µl, 1 µl, and 2 µl, respectively. The PCR program was exactly the same.

3.1.5.3 – Gel purification of Bartram products

The primers and primer dimers of Bartram products were cleaned by running the samples through a 2% metaphor agarose gel (Lonza group, Basel, Switzerland), followed by a purification using the NucleoSpin Gel and PCR Clean-up Kit (Macherey-Nagel, Düren, Germany), according to the instructions of the manufacturer. The purified products were then quantified with the Qubit® dsDNA HS Assay Kit (Life Technologies, Carlsbad, CA, US) before sequencing.

3.1.6 – Sequencing with Illumina

Approximately 53 samples were pooled together, each with the same concentration. The first pooled sample was checked with a 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, U.S). The samples from the first scoping study (collected from Halle) were sequenced on the Illumina HiSeq 2500 platform (Illumina, San Diego, CA, USA), whereas samples from the second scoping study and the full-factorial crossed design experiment (collected from Berlin) were sequenced on the Illumina Nextseq 500 platform (Illumina, San Diego, CA, USA), with the employment of a paired-end run.

3.1.7 – Downstream sequencing data processing

3.1.7.1 – Pre-processing raw sequence reads with Qiime2

The QIIME2™ plugins (version 2018.11; https://qiime2.org/) were utilized to assess the raw sequence reads of Illumina (Caporaso *et al.* 2010). The sequences were demultiplexed, imported to qiime2-artifacts with “tools import”, joined with “vsearch”, quality-filtered with “quality filter”, and denoised with “deblur”. All plugins were employed with the default options (Bokulich et al., 2013; Amir et al., 2017) and deblur was applied with a trim length of 165 bases. Furthermore, *de novo* multiple sequence alignment was done using “alignment” plugin with “mafft“ as parameter. The un-conserved regions and columns with high number of gaps were masked with “mask” plugin. A midpoint rooted tree was then constructed using “phylogeny” plugin with “fastree” as parameter. The SILVA database (version 138), which was trained on the V3 region of 16S ribosomal RNA gene sequence, was employed as the taxonomic classifier of the representative sequences. The taxonomic classification was performed with “feature-classifier” plugin. The read count table (table.qza), the taxonomic table (table.qza), and the rooted-phylogenetic tree (rooted-tree.qza) were retrieved and converted accordingly. The outputs were then imported to R software for deeper analysis with the R package phyloseq (McMurdie and Holmes, 2013) where sequences associated to chloroplast and mitochondria were removed in prior.

3.1.7.2 – Statistical analysis

All statistical analysis was performed in R program (version 3.6.1; R Core Team 2019). The rarefaction analysis, including rarefaction to an even depth, was performed using the R package phyloseq (McMurdie & Holmes 2013). Sample coverages estimates (estimation of the portion of population that is covered by the sample) were calculated with the iNEXT package (Chao et al. 2014; Hsieh et al. 2020). The α-diversity estimates (richness, Shannon, and evenness) were analysed in detail using the R package RAM (Chen *et al.* 2018). The alpha value of gambin distribution was calculated with the R package gambin to analyse the species abundance distribution (SAD) across samples where low values suggest logseries-like SAD (the community members belong to a few species and each species is represented by very few individuals) and high values suggest lognormal-like SAD (only few species have either high or low abundances while most have intermediate abundance) (Matthews *et al.* 2014; de Miranda *et al.* 2019). Therefore, the alpha gambin is advantageous to compare different shape of SAD across distinct samples. In addition, while evenness is dependent on species richness, this is not the case for alpha gambin. Thus, alpha gambin fits better compared to evenness to evaluate species distribution across samples with distinct species richness. The visualization of alpha diversity and the subsequent plots was done using the R package ggplot2, unless mentioned otherwise (Wickham *et al.* 2016).

For investigation of differences in bacterial communities, the distance based on weighted UniFrac (Lozupone *et al.* 2007) was calculated using the phyloseq package (McMurdie & Holmes 2013) and employed to construct the Non-Metric Multidimensional Scaling (NMDS) plot. NMDS was performed in 3 dimensions (k = 3) although sometimes the visualization was only depicted in 2 dimensions. Only ordination that resulted in stress values above 0.11 was visualized since it is considered a fair and good fit ordination that is not based on random choices. The significant differences were tested by permutational analysis of variance (PERMANOVA; 999 permutations) with the R package vegan (Oksanen *et al.* 2019).

The contribution of different variables to the bacterial community composition was determined with variance partitioning analysis, also available in vegan (Oksanen *et al.* 2019). Apart from the plots constructed using ggplot2, bacterial community composition was also visualized using R metacoder package (Foster *et al.* 2017). Unlike ggplot2 bar plot that only depicted a specific taxonomic rank, metacoder provides a tree-based visualization which displays all taxonomic ranks from phylum to genus level in a single plot. In addition, metacoder was also used to perform pairwise comparison of bacterial communities between two distinct community type (total and active bacterial community that were based on DNA and RNA data, respectively), habitat types (bulk soil vs rhizosphere), soil

substrates (alkaline vs acidic), and distinct plant species (all possible pairwise combination between *H. perforatum, H. olympicum*, and *H. balearicum*). The results were visualized as phylogenetic trees where enriched taxa of each of the variables (based on Wilcoxon rank-sum test and followed by a Benjamini-Hochberg (FDR) correction for multiple comparisons) were colored differently. Moreover, visualization of active taxa of specific group (either bulk soil or the rhizosphere or inside roots of a particular plant species) and in a specific soil substrate (either acidic or alkaline, specifically for the full-factorial crossed design study) was also conducted with metacoder. The resulted plots comprising phylogenetic trees (one tree for each single group) where each node (represents different taxonomic rank) was colored in different shades of green depending on how many taxa (at sequence variant level) were classified into that rank. For each of the comparisons, the taxonomy detail of each node was given in a key tree.

The significant differences in the relative abundance of bacterial taxa, alpha diversity measures, percentage of active taxa, and plant biomass were tested with a multiple comparison analysis of variance using Tukey's all pair comparisons with the R multcomp package (Hothorn *et al.* 2008). The t-test was employed to compare two distinct groups using the R stats package (R Core Team 2019). A pairwise comparison between two factors, to determine significantly enriched taxa, was performed with the R package metacoder, applying the Wilcoxon rank-sum test. The p values were adjusted with the Benjamini-Hochberg (FDR) correction for multiple comparisons (Foster *et al.* 2017) when distinct nucleic acid data (DNA vs RNA) and soil substrates (acidic vs alkaline) were compared, independently of the habitat types and plant species.

The active taxa were identified at the sequence variant level using a specific threshold (see the next subchapter for the detail). The ggtree (Yu *et al.* 2017) and ggplot2 (Wickham *et al.* 2016) packages were employed to construct the phylogenetic trees of the plant-species specific active taxa along with the corresponding heatmaps. To determine if the plant species-specific active taxa, either in the rhizosphere or inside roots, were phylogenetically clustered at deep or terminal branches when compared to a null model, the net relatedness index (NRI) and nearest taxon index (NTI) were calculated using ses.mpd and ses.mntd functions available in the R package picante (Kembel *et al.* 2010).

3.1.7.3 – Determination of a specific threshold based on rRNA:rDNA ratio to identify active taxa

The threshold for the identification of active taxa was developed by Johannes Sikorski of DSMZ. The rRNA:rDNA ratio has been used previously to define the metabolic capacity of microbial cells *in situ*. The utilization of the rRNA:rDNA ratio has revealed that a small portion of the community, with low

abundance (rare taxa), contains remarkably high rRNA:rDNA ratio values and thus often referred as active. To determine if the pattern is a truly biological one and not an artefact, an investigation of two African soil samples (KS and OFP) was conducted. For each soil sample, 4 parallel RNA extractions were employed (labeled as A, B, C, and D) and for each single RNA extract, 4 parallel Bartram PCRs were performed (labeled as 1, 2, 3, and 4). These results in 16 technical replicates for each soil sample that were subsequently sequenced on the Illumina Hiseq and Nextseq platforms. In total, there were 64 individual replicates (4x16; 4 datasets with different combinations of soil samples and sequencing machines, each with 16 technical replicates). Each dataset was analysed individually at sequence variant level. For each dataset, the 4 PCR replicates (1, 2, 3, and 4) from a specific RNA extract (A/B/C/D) were pairwise-compared with another 4 PCR replicates belonging to a different batch of the RNA extraction (for example A1, A2, A3, and A4 were each individually compared to B1, B2, B3, and B4, that result in 16 different combinations). Since there were 4 batches of RNA extraction, we ended up with 96 distinct comparisons for each dataset (6 different combinations of RNA extracts (A-B, A-C, A-D, B-C, and B-D) multiply by 16).

For each dataset, the 96 unique combinations were visualized together where the taxa abundance in pair number 1 was plotted against the abundance in pair number 2. In addition, additional plot was constructed where x axis depicts the abundance of taxa in pair number 1 and y axis depicts the ratio of abundance between pair number 2 and 1 (Figure 9). Only taxa that at least present in one of the pairs were investigated. The pseudocount of 1 was introduced on the x axis (labeled as ‘1’ in the plot) for the taxa that only present in one of the pairs. This specific plot has been used to investigate the metabolic capacity of microbial cells based on the rRNA:rDNA ratio values, where pair number 2 represents the active (RNA-based) community and pair number 1 represents the total (DNA-based) community. The rRNA:rDNA ratio has been reported to decline along the x axis (with increasing DNA abundance). In addition, the majority of soil taxa has low DNA abundance and thus the number of taxa also declines along the x axis.

As each pair of all 96 combinations were actually a technical replicate of the same exact soil sample, the abundance of taxa would ideally be the same between the two, and thus the ratio (y axis) would be constantly 1. However, we observed taxa with the sample2:sample1 ratio higher than 1, implying that even stochastic variation associated with different batches of RNA extraction, PCR, and to some extent sequencing artefacts could lead to the false positive “active” taxa (Figure 9). The same pattern was observed regardless of the soil samples (KS and OPF) and the sequencing platforms.

A curve fitted to the maximum values of the sample2:sample1 ratio along the x axis was subsequently plotted. The exact same curve was then re-drawn on the real plot of the rRNA:rDNA

ratio belonging to 60 different grasslands samples. By visual inspection, more taxa on the real rRNA:rDNA plot have the ratio above the curve compared to those resulted from the stochastic variation of the read counts. This result implies that the rate of false positive active taxa is minimized with increasing values of rRNA:rDNA ratio. However, the curve that was fitted to the maximum values is also not ideal since it increases the rate of false negative active taxa that were plotted below the curve. For this reason, we developed an optimal threshold that compromises a balance between the occurrences of false positive and false negative active taxa. Nonetheless, a validation based on the distribution of active taxa across biological replicates would be done subsequently to minimize the false positive active taxa (see the next subchapter).

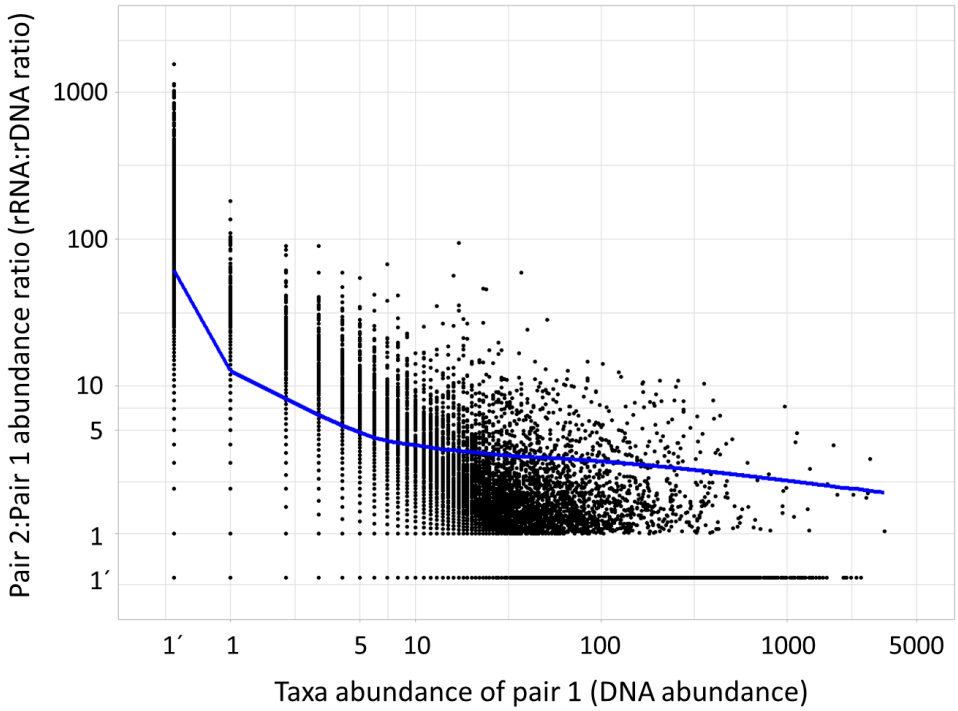


Figure 9 – The illustration of pairwise comparison of sequences belonging to different batches of RNA extraction (4 batches in total, each amplified in 4 distinct PCR reactions) for each soil sample (either KS or OFP) that was sequenced in a specific sequencing machine (either HiSeq or Nextseq). Altogether, 96 unique combinations (6 unique combination of RNA extracts (A-B, A-C, A-D, B-C, B-D, C-D) x 16 PCR combinations) were available for each soil sample (dependent of the sequencing machine) and the data were merged together in a single plot. The pseudocount of 1 was introduced on the x axis (labeled as '1' in the plot) for the taxa that only present in one of the pairs. The same plot has been used to investigate the activity of bacteria based on rRNA:rDNA ratio where pair no. 1 represents DNA abundance and pair no. 2 represents RNA abundance. Blue line depicts the threshold to identify active taxa that cannot be explained by stochastic variation in the sequence read counts.

To determine the specific threshold, 200 random samples were generated for each x axis value of the technical replicates plot and only the maximum value was stored for each x axis. The random sampling was performed 500 times, that results in 500 different maximum values for each x axis. Since the number of taxa decreases along the x axis, many high x axis values contain less than 200 taxa. For this reason, the "sliding window" was applied to ensure each sampling point has at least 200 taxa for the random sampling and started at the x axis value of 10. For values between 10 and 300, a 10% window in both directions was applied. For example, the x axis value of 13 will results in the window values between 11.7 and 14.7, and thus the x axis values of 12, 13, and 14 were merged together. Furthermore, a 30% window was applied for x axis values of 300 to 1049 whereas all values between 1050 and 1950 were merged together as a single sampling point (x axis = 1500). For each of these sampling points, a median value was calculated from the 500 different maximum values. The threshold was then created by fitting these median values into a curve along the x axis. In order to obtain a smooth threshold-curve, the partial loess and polynomial spline fitting was applied (Figure 9; blue line represents the threshold). The active taxa were defined as the ones with the rRNA:rDNA ratio above the threshold. At the end, the rate of false positive active taxa based on this threshold is very low (0.4%).

3.1.7.4 – Validation of the threshold based on the distribution of active taxa across biological replicates

The threshold to identify active bacteria can be considered as effective if it is successfully identified bacteria that are active in multiple biological replicates, suggesting that the identification is not based on random processes. Thus, the threshold should not be utilized further if it is only identified taxa that are only active in 1 biological replicate. In order to validate the ability of the threshold to detect the same active taxa across biological replicates and further minimize the rate of false positive active taxa, the distribution of active taxa across biological replicates of each group (bulk soil/rhizosphere/roots of a particular plant species) was investigated. The result was then compared to a stochastic null model obtained from a random sampling. We hypothesize that the threshold will identify the same active taxa in higher number of biological replicates compared to the output of random sampling, implying the reliability of the threshold to identify the real active taxa.

For each group, the active taxa were sorted based on the occurrences across biological replicates (detected to be active only in one or up to 6 biological replicates) and the percentage of the active taxa for each number of biological replicates was calculated. For each individual biological replicate, random sampling of sequence variants was conducted with the sampling size equal to the number of active taxa. The randomly-sampled sequence variants were also sorted based on the occurrences

across biological replicates and the percentage for each number of replicates was calculated. This process was repeated 1000 times. Both the percentage of active and randomly-sampled taxa were visualized in one plot where x axis depicts the number of biological replicates and y axis depicts the percentage of the active or randomly-sampled taxa. The percentage of active taxa was displayed as a stacked bar whereas the percentage of randomly-sampled ones with 1000 replications was displayed as a boxplot. Of randomly-sampled active taxa, only values above 0 were stored out of 1000 repetitions for higher x axis values (x axis > 2), as it is likely that the majority of the randomly-sampled taxa occur only in 1 and 2 biological replicates and thus the boxplots with higher x axis values will be positioned almost at the bottom of y axis. Furthermore, taxonomic affiliation of the active taxa was visualized with the R package metacoder (Foster *et al.* 2017).

3.2 – Cultivation of phylogenetically novel bacterial taxa that might elicit hypericin and hyperforin

The functional roles and ecological relevance of soil bacterial community has not been fully understood as the majority of taxa still resist cultivation. The phylogenetically novel taxa are likely performed ecological functions that have not been deeply investigated. This may be related to the production of hypericin and hyperforin for bacteria associated with *Hypericum*. Therefore, cultivation was performed to access potential taxa that might elicit hypericin and hyperforin.

3.2.1 – Soil samples for cultivation

High throughput cultivation was conducted at the end of 2016 after the first sampling campaign (the first scoping study) in Leibniz IPB Halle was executed. The following studies (the second scoping study and the full-factorial crossed design experiment) were not assisted by cultivation attempts due to the fact that the cultivation is a highly time-consuming process (Overmann 2013). In total, 6 samples collected from Halle were selected for high throughput cultivation, each representing a single group of bulk soil/rhizosphere of a particular plant species except for the rhizosphere of field *H. perforatum* that was represented by 2 soil samples since the plant species is the most well-known producer of hypericin and hyperforin. Less samples of *H. polyphyllum* were selected as the content of hypericin and hyperforin had been reported to be lower for the plant species compared to *H. perforatum*. Specifically, the samples included bulk soil and the rhizosphere of wild *H. perforatum* (one bulk soil and two rhizosphere samples) and *H. polyphyllum* (one sample each), and one rhizosphere sample of the greenhouse-cultivated *H. perforatum*. The pH was measured in water and 2 mM $CaCl_2$ and employed as the pH of the cultivation media.

3.2.2 – High throughput cultivation in liquid media

Total bacterial cell numbers were determined using the SYBR Green I staining (Life Technologies) as described earlier (Lunau *et al.* 2005). The soil was suspended in HEPES buffer (10 mM, pH 7) and 20 µl of this suspension, with either 25 or 50 cells, was transferred to 96-well microtiter plates containing 180 µl of either SSE/HD 1:10 or R2A1:10 medium buffered at pH 7. The SSE/HD1:10 is a growth medium consisting of soil solution equivalent (Angle *et al.* 1991) with 10- vitamin (Balch *et al.* 1979) and trace element SL-10 (Tschech & Pfennig 1984) solutions, buffered with HEPES at pH 7 and amended by peptone, yeast extract and glucose (Huber *et al.* 2016; Vieira *et al.* 2017). The outer wells of the microtiter plates were left un-inoculated and the plates were incubated for 3 months at room temperature.

Growth was detected by turbidity and the grown wells were selected for analysis of the 16S rRNA V1-V2 hypervariable regions, using a barcoded Illumina paired-end sequencing method (Camarinha-Silva *et al.* 2014) with modified PCRs. The V1-V2 regions were selected since the sequence length is longer compared to the V3 region and thus allows more accurate taxonomic assignment. The freeze and thaw cycles were performed on each grown well, where the cells were disrupted by incubation in 70% ice-cold EtOH for 5 min, followed by incubation at 99 °C for 5 min using the Thermoblock.

The first PCR reaction (pre-amplification PCR) was performed to amplify the V1-V2 regions of 16S rRNA using primer pair 27F (5'-AGAGTTTGATCCTGGCTCAG -3') and 338R (5'-GCTGCCTCCCGTAGGAGT -3'). The amplification was conducted in a total volume of 20 µL consisted of 4.0 µL PCR buffer (5x, including MgCl2), 0.4 µL dNTPs (10 mM each), 0.5 µL each of forward and reverse primers (10 pmol µL-1), 0.2 µL Phusion Taq DNA polymerase (2U μL^{-1} Thermo Scientific) and 1.0 µL the DNA template. The PCR program was started with an initial denaturation step at 98 °C for 30 min, followed by 25x cycles of denaturation at 98 °C for 10 s, annealing at 55 °C for 15s, and elongation at 72 °C for 30 s, and finalize with an elongation step at 72 °C for 7 min.

The second PCR was conducted to incorporate barcodes and the Illumina sequencing primers (ISP) that were attached to the forward primer. The product of the first PCR was employed as the template. The PCR mixed was similar to the first one except for the volume of dNTPs (1.6 µL). An initial denaturation step was employed at 98 °C for 30 s, followed by 15 cycles of denaturation at 98 °C for 10 s, annealing at 55 °C for 10 s, and elongation at 72 °C for 45 s. The program was finalized with an elongation step at 72 °C for 7 min.

The third PCR was performed to integrate Index and the Illumina adaptors (IAdapt). The PCR reaction was performed in a total volume of 30 µL, which consist of 6.0 µL Buffer (5x, including

MgCl2), 2.4 µL dNTPs (10 mM each), 0.3 µL Phusion Taq DNA polymerase (2U μL^{-1} Thermo Scientific), 0.6 µL of the second PCR product, and 0.75 µL each of multiplexing sequencing forward primer and index reverse primer (10 pmol μL^{-1}, both contained Illumina Adaptors). The PCR program was the same as the second PCR. The sequencing was performed on the Illumina MiSeq platform (Illumina, San Diego, CA, USA).

The bioinformatic analysis was performed as described previously (sub-chapter 3.1.7.1), with a trim length of 276 bp employed for deblur. The taxonomy of the reads was assigned to the silva database (v.138) (Quast *et al.* 2013) with uclust (Edgar 2010). The cultures with sequences sharing less than 97% similarity to validly described strains were purified on SSE/HD1:10 or R2A1:10 medium buffered at pH 7 and solidified with 0.8% (w/v) gellan gum.

3.2.3 – PCR identification for purified strains

The colony PCR was employed to identify purified strains. The almost full-length 16S rRNA gene sequences were amplified using primer forward 8f (5′-AGAGTTTGATCCTGGCTCAG-3′) and primer reverse 1492r (5′-GGTTACCTTGTTACGACTT-3′) (Turner *et* al. 1999; Lane 1991). The PCR was performed in a total volume of 20 µL, which consist of 2.0 µL PCR buffer (10x), 0.4 µL dNTP mix (10 mM each), 0.08 µL of each primer pair (50 pmol μL^{-1}), 0.8 µL $MgCl_2$ (25 mM), 0.4 µL BSA (20 mg mL^{-1}), and a tiny dab of a pure colony. The PCR program was started with an initial denaturation step at 94 °C for 10 min, followed by 30 cycles of denaturation at 94 °C for 30 s, annealing at 56 °C for 30 s, and elongation at 72 °C for 1 min, and ended with a final elongation step at 72 °C for 7 min. The Sanger sequencing was performed using the AB 3730 DNA analyzer (Applied Biosystems, Foster City, CA) by employing AmpliTaq® FS BigDye® Terminator Cycle Sequencing Kit. The primer pairs 1055f (5′- ATGGCTGTCGTCAGCT-3′) (Lane 1991) and 341r (5′-CTGCTGCCTCCCGTAGG-3′) (Muyzer et al. 1993) was used as the sequencing primers and the output sequences were aligned to the EzBioCloud database (Yoon *et al*. 2017)

3.2.4 – Co-cultivation test of Acidobacteriota strain

The experiment was conducted to determine the effect of adding helper bacteria supernatant to the growth of an Acidobacteriota strain that cannot be grown axenically. The strain is a novel species candidate of the genus *Stenotrophobacter. Pseudomonas prosekii* is one of the accompanied strains inside the microtiter well and thus selected as the helper bacterium. The helper bacterium was inoculated to the liquid LB medium and incubated for 4 days at 28 °C. The supernatant was separated from the cells through centrifugation and filtration using a 0.22 µm filter (Merck Millipore). A total of 6 tubes were prepared with each contains 0.1, 0.25, 0.5, 1, 2, and 2.5 mL of

helper bacteria supernatant. A fresh liquid SSE/HD1:10 medium was added to the tubes to a total volume of 4.5 mL. A 0.5 mL of the Acidobacteriota inoculum was then added and the tubes were incubated aerobically for approximately a month. Growth was monitored by measuring OD_{660} every day. In addition, a positive control (inoculated with both target strain and the helper bacterium) and two negative controls (the target strain was grown alone in either SSE/HD1:10 medium or with the addition of 1 mL fresh LB medium) were prepared.

3.2.5 – Characterization of *Hypericibacter terrae* R5913^T and *Hypericibacter adhaerens* R5959^T

Two new species of a novel genus of the family *Rhodospirillaceae* were successfully isolated and fully characterized, for which the names *Hypericibacter terrae* sp. nov. (type strain = R5913^T) and *Hypericibacter adhaerens* sp. nov. (type strain = R5959^T) were proposed.

3.2.5 – Phenotypic and chemotaxonomic characterization of the novel strains

Phenotypic characterization was conducted according to the methods described earlier (Huber *et al.* 2014; Pascual *et al.* 2015a; Pascual *et al.* 2015b). Cells were observed under the light microscope Zeiss Axio Lab. A1 (Carl Zeiss, with AxioCam Mrm camera) after Gram and malachite green staining. For scanning electron microscopy, cells in liquid R2A medium were fixed in 2% glutaraldehyde for 30 min at room temperature, mixed further with 5% paraformaldehyde and stored at 4 °C. The cells were subsequently washed with TE buffer (10 mM Tris, 2 mM EDTA, pH 6.9) and dehydrated on ice in a graded series of acetone (10, 30, 50, 70, 90 and 100%; each for 10 min). Samples were subjected to a critical-point drying by applying CO_2 (CPD 030; Bal-Tec). The dried samples were coated with a gold/palladium (80:20) film by sputter coating (SCD 500; Bal- Tec) before examination with the Zeiss Merlin field emission scanning electron microscope using the Everhart Thornley HESE2- detector and the inlens SE- detector in a 25:75 ratio at an acceleration voltage of 5 kV. Images were recorded with the Zeiss SEMSmart V 5.05 and the contrast and brightness were adjusted with Adobe Photoshop CS5.

Catalase activity was assessed by observing gas formation after cell exposure to 3% (v/v) H_2O_2 (Taylor & Achanzar 1972) whereas cytochrome-c oxidase activity was determined using the Bactident Oxidase (Merck), following the instructions of the manufacturer. The growth ranges and optima for temperature, pH and salinity were determined aerobically in triplicates in the liquid R2A medium. Growth was tested across the temperature range of 10–45 °C and pH optima were determined between pH values of 1.0 and 11.0 as described earlier (Huber *et al.* 2014; Huber *et al.* 2016; Vieira *et al.* 2016; Foesel *et al.* 2013). Depending on the pH, MES, HEPES, HEPPS or CHES (Sigma- Aldrich or AppliChem; 10 mM) were utilized as buffers. The salt tolerance was determined in

the liquid R2A medium supplemented with 0, 0.25, 0.5, 1, 3, 5, 7.5 and 10% (w/v) NaCl. The growth was monitored by measuring the OD_{660} and the optimal growth was defined as ≥ 75% of the highest growth rate observed.

The growth of strains R5913^T and R5959^T under microaerophilic and anaerobic conditions was assessed in a candle jar (Gerhardt *et al.* 1981) and using Anaerocult P (Merck), respectively. Exoenzyme activities, nitrate reduction, indole production, fermentation of glucose, arginine dihydrolase, urease and β-galactosidase (PNPG) activities, and the hydrolysis of aesculin and gelatin were tested using the commercial API 20NE and API ZYM galleries (bioMérieux), according to the instructions of the manufacturer. The assimilation of different carbon substrates in the API 20NE was not conducted because the medium was not suitable for the growth of R5913^T or R5959^T.

The ability of strains R5913^T and R5959^T to utilize 108 different carbon substrates, including various sugars, organic acids, keto acids, alcohols, amino acids, casamino acid, casein hydrolysate, laminarin, peptone, yeast extract and Tween 80, was tested in triplicate using liquid R2A medium. The final concentrations were as described previously (Huber *et al.* 2014; Huber *et al.* 2016). A positive outcome was defined when the mean of OD_{660} of the three parallel tests was 1.5 times or higher than that of the control (culture without substrate). The weak positive growth was defined as reaching values between 1.2 and 1.5 times the controls. Additionally, the abilities of both strains to utilize complex carbon substrates such as starch, cellulose, carboxymethyl (CM)-cellulose, chitin, lignin, polygalacturonic acid, pectin, xylan and 2,2'-azinobis (3- ethylbenzthiazoline-6- sulfonate) (ABTS) were tested on the solidified R2A medium. Specific staining solutions were employed, as described previously (Pascual *et al*. 2015a), to detect degradation of the polymeric substrates after approximately 1 month of incubation.

Strains R5913^T and R5959^T were harvested during the late logarithmic phase for the detection of fatty acids, polar lipids and respiratory quinones. Fatty acids were extracted, saponified and methylated using the standard protocols and identified with the TSBA40 library (MIDI Microbial Identification System; version 6.1 (Sasser 2001)). Polar lipids were assessed by two-dimensional TLC (Bligh & Dyer 1959; Tindall *et al.* 2007). Isoprenoid quinones were extracted from dried biomass using chloroform/methanol solution (2:1, v/v) (Collins & Jones 1981) and analysed further with HPLC (Tindall 1990).

The almost full-length 16S rRNA gene sequences of strains R5913^T and R5959^T (both 1441 bp) were amplified and sequenced with the method described earlier (Pascual *et al.* 2015b). Multiple sequence alignment was performed using the sina alignment tool available in the arb-silva website

(Pruesse *et al.* 2012) and the result was visually inspected to re-evaluate and improve the uncertain alignments. Phylogenetic analysis was performed on mega (version 7.0; Kumar *et al.* 2016), involving the reconstruction of neighbour-joining (NJ; with Kimura's two-parameter evolutionary model) and maximum-likelihood (ML; T92+I+G evolutionary model) trees. Tree topology was evaluated by bootstrap analysis using 1000 replications. To determine the abundance of both novel strains in similar habitats, the almost-full length 16S rRNA gene sequences were aligned at 99% similarity to the databases in the Integrated Microbial Next Generation Sequencing (IMNGS) depository (Lagkouvardos et al. 2016).

For genome sequencing, the DNA was isolated using the Qiagen Genomic-tip 100/G (Qiagen) according to the manufacturer instructions. An SMRTbell template library was prepared according to the PacificBio-sciences instructions, employing the 'Procedure and Checklist – Greater Than 10 kb Template Preparation'. An 8 µg of genomic DNA was sheared briefly for the preparation of 15 kb libraries using g-tubes from Covaris, following the instructions of the manufacturer. The DNA was end-repaired and ligated overnight to the hairpin adapters, applying components from the DNA/Polymerase Binding Kit P6 from Pacific Biosciences. The reactions were carried out according to the manufacturer instructions. The BluePippin size-selection of fragments greater than 4 kb was performed following the instructions of the manufacturer (Sage Science). The conditions for annealing of sequencing primers and the binding of polymerase to purified SMRTbell template were assessed with the Calculator in RS Remote (Pacific Biosciences). One SMRT cell was sequenced on the PacBio RSII platform (Pacific Biosciences), taking one 240 min movie for each strain. Libraries for sequencing on the Illumina platform were prepared employing the Nextera XT DNA Library Preparation Kit (Illumina), with some modifications according to Baym *et al.* (Baym *et al.* 2015). Samples were sequenced on the NextSeq 500 sequencer. Genome assembly was performed applying the RS_HGAP_Assembly.3 protocol that was included in SMRT Portal version 2.3.0. Error-correction was performed by mapping of the Illumina short reads onto the finished genomes using the Burrows–Wheeler Aligner bwa 0.6.2 in paired-end (sampe) mode using default settings (Li & Durbin 2009) with subsequent variant and consensus calling with VarScan 2.3.6 (Koboldt *et al.* 2012). In addition, automated genome annotation was carried out using dfast (Tanizawa *et al.* 2018). The visualization of both genomes was done with artemis (Carver *et al.* 2012).

The average nucleotide identity (ANI) value was calculated between the two strains, and with their phylogenetically closest neighbors, using the OrthoANIu algorithm available in the EzGenome web service (www.ezbiocloud.net/tools/ani) (Yoon *et al.* 2017). A phylogenomic tree based on nucleotide sequences was generated with the aim of obtaining more accurate phylogenetic inferences of both

strains. The UBCG v. 3.0 pipeline (up-to-date bacterial core gene set) (Na *et al.* 2018) was utilized to reconstruct an ML tree based on a multiple alignment of a set of 92 universal and single copy gene sequences using the tool FastTree v2,10,1.

3.2.6 – Statistical analysis

Pre-processing of high throughput cultivation data was performed using plugins available in the QIIME2™ (version 2018.11; https://qiime2.org/) (Caporaso *et al.* 2010), with the same protocols as described for the bacterial community data. The statistical analysis was conducted in R (version 3.6.1; R Core Team 2019) with rarefaction to the same number of reads per sample was performed in prior with the phyloseq package (McMurdie & Holmes 2013). Unless mentioned otherwise, visualization was performed with the ggplot2 package (Wickham *et al.* 2016). The iNEXT package (Chao et al. 2014; Hsieh et al. 2020) was utilized to calculate sample coverage estimates (estimation of the portion of population that is covered by the sample). Alpha diversity measures (richness, Shannon, evenness, and alpha gambin) were calculated with the RAM package (Chen *et al.* 2018). The significant difference in alpha diversity measures was tested with the multiple comparison analysis of variance using Tukey with the R multcomp package (Hothorn *et al.* 2008). T-test was employed for specific analysis where two groups were compared, using the stats package (R Core Team 2019). The ordinations were calculated using Non-Metric Multidimensional Scaling (NMDS) (R vegan package; Oksanen *et al.* 2019), based on weighted UniFrac distances (R phyloseq package; McMurdie & Holmes 2013). NMDS was performed in 3 dimensions (k = 3). Only ordination that resulted in stress values above 0.11 was visualized since it is considered a fair and good fit ordination that is not based on random choices. Taxonomic composition and taxa enrichment based on Wilcoxon rank-sum test between two different growth media were analysed with the R metacoder package (Foster *et al.* 2017).

In order to determine if some bacterial taxa are consistently isolated together during cultivation than expected under a null model, co-occurrence test based on relative abundance of bacterial taxa growing (abundance> 1%) inside the wells of microtiter plates was investigated using the make_netassoc_network function available in R netassoc package (Morueta-Holme & Blonder 2016) with default options except for the alpha (0.01 instead of 0.05). The null model of community assembly was calculated using the permatfull function available in R vegan package (Oksanen *et al.* 2019). The data was visualized with the R igraph package (Csardi & Nepusz 2006).

3.3 – References

Angle, J. S., Mcgrath, S. P. & Chaney, R. L. 1991. New culture medium containing ionic concentrations of nutrients similar to concentrations found in the soil solution. *Applied and Environmental Microbiology*, **57**, 3674-3676.

Amir, A., McDonald, D., Navas-Molina, J. A., Kopylova, E., Morton, J. T., Xu, Z. Z., Kightley, E. P., Thompson, L. R., Hyde, E. R., Gonzales, A. & Knight, R. 2017. Deblur rapidly resolves single-nucleotide community sequence patterns. *mSystems,* **2**, doi: 10.1128/mSystems.00191-16.

Baym, M., Kryazhimskiy, S., Lieberman, T. D., Chung, H., Desai, M. M. & Kishony, R. 2015. Inexpensive multiplexed library preparation for megabase-sized genomes. *PLoS One*, **10**, doi: 10.1371/journal.pone.0128036.

Balch, W. E., Fox, G. E., Magrum, L. J., Woese, C. R. & Wolfe, R. S. 1979. Methanogens: reevaluation of a unique biological group. *Microbiological Reviews*, **43**, 260-296.

Barillot, C. D. C., Sarde, C., Bert, V., Tarnaud, E. & Cochet, N. 2013. A standardized method for the sampling of rhizosphere and rhizoplan soil bacteria associated to a herbaceous root system. *Annals of Microbiol*ogy, **63,** 471-476, doi: 10.1007/s13213-012-0491-y.

Bartram, A. K., Lynch, M. D. J., Stearns, J. C., Moreno-Hagelsieb, G. & Neufeld, J. D. 2011. Generation of multimillion-sequence 16S rRNA gene libraries from complex microbial communities by assembling paired-end Illumina reads. *Applied and Environmental Microbiology*, **77**, 3846-3852, doi: 10.1128/AEM.02772-10.

Bligh, E. G. & Dyer, W. J. 1959. A rapid method of total lipid extraction and purification. *Canadian Journal of Biochemistry and Physiology*, **37**, 911-917.

Morueta-Holme, N., Blonder, B., Sandel, B., McGill, B., Peet, R., Ott, J. E., Violle, C., Enquist, B., Jørgensen, P., & Svenning, J. 2016. A network approach for inferring species associations from co-occurrence data. *Ecography*, **39**, 1139-1150, doi: 10.1111/ECOG.01892.

Bokulich, N. A., Subramanian, S., Faith, J. J., Gevers, D., Gordon, J. I., Knight, R., Mills, D. A. & Caporaso, J. G. 2013. Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. *Nature Methods*, **10**, 57-59, doi: 10.1038/nmeth.2276.

Camarinha-Silva, A., Jáuregui, R., Chaves-Moreno, D., Oxley, A. P. A., Schaumburg, F., Becker, K., Wos-Oxley, M. L. & Pieper, D. H. 2014. Comparing the anterior nare bacterial community of two discrete human populations using Illumina amplicon sequencing. *Environmental Microbiology*, **16**, 2939-2952, doi: 10.1111/1462-2920.12362.

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., Fierer, N., Pẽa, A. G., Goodrich, J. K., Gordon, J. I., Huttley, G. A., Kelley, S. T., Knights, D., Koenig, J. E., Ley, R. E., Lozupone, C. A., McDonald, D., Muegge, B. D., Pirrung, M., Reeder, J., Sevinsky, J. R., Turnbaugh, P. J., Walters, W. A., Widmann, J., Yatsunenko, T., Zaneveld, J. & Knight, R. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nature Methods*, **7**, 335–336, doi: 10.1038/nmeth.f.303.

Carver, T., Harris, S. R., Berriman, M., Parkhill, J. & McQuillan, J. A. 2012. **Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data.** *Bioinformatics, **28**, 464-469, doi*: 10.1093/bioinformatics/btr703.

Chao, A., Gotelli, N. J., Hsieh, T. C., Sander, E. L., Ma, K. H., Colwell, R. K. & Ellison, A. M. 2014. Rarefaction and extrapolation with Hill numbers: A framework for sampling and estimation in species diversity studies. *Ecological Monographs*, **84**, 45-67, doi: 10.1890/13-0133.1.

Chen, W., Simpson, J. &, Levesque, C. A. 2018. RAM: R for Amplicon-Sequencing-Based Microbial-Ecology. *R package version 1.2.1.7*, https://CRAN.R-project.org/package=RAM.

Collins M.D. & Jones, D. 1981. Distribution of isoprenoid quinone structural types in bacteria and their taxonomic implication. *Microbiology Reviews*, **45**, 316-354.

Csardi, G. & Nepusz, T. 2006. The igraph software package for complex network research. *InterJournal Complex Systems*, 1695, doi: 10.1177/001316446902900315.

de Miranda, M. D., Borda-de-Água, L., Pereira, H. M. & Merckx, T. 2019. Species traits shape the relationship between local and regional species abundance distributions. *Ecosphere*, **10**, doi: 10.1002/ecs2.2750.

Edgar, R. C. 2010. Search and clustering orders of magnitude faster than blast. *Bioinformatics,* **26**, 2460-2461, doi: 10.1093/bioinformatics/btq461.

Foesel, B. U., Rohde, M. & Overmann, J. 2013. Blastocatella fastidiosa gen. nov., sp. nov., isolated from semiarid savanna soil - the first described species of Acidobacteria subdivision 4. *Systematic and Applied Microbiology*, **36**, 82-89, doi: 10.1016/j.syapm.2012.11.002.

Foster, Z. S. L., Sharpton, T. J. & Grünwald, N. J. 2017. Metacoder: An R package for visualization and manipulation of community taxonomic diversity data*. PLOS Computational Biology*, **13**, doi: 10.1371/journal.pcbi.1005404.

Gerhardt, P., Murray, R. G. E., Costilow, R. N., Nester, E. W. & Wood, W. A. (eds). 1981. *Manual of Methods for General Bacteriology*. American Society for Microbiology, Washington, DC.

Hothorn, T., Bretz, F. & Westfall, P. 2008. Simultaneous inference in general parametric models. *Biometrical Journal*, **50**, 346-363, doi: 10.1002/bimj.200810425.

Hsieh, T. C., Ma, K. H. & Chao, A. 2016. iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers). *Methods in Ecology and Evolution*, **7**, 1451-1456, doi: 10.1111/2041-210X.12613.

Huber, K.J., Geppert, A. M., Wanner, G., Fösel, B. U., Wüst, P. K. & Overmann, J. 2016. The first representative of the globally widespread subdivision 6 Acidobacteria, Vicinamibacter silvestris gen. nov., sp. nov., isolated from subtropical savannah soil. *International Journal of Systematic and Evolutionary Microbiology*, **66**, 2971–2979, doi: 10.1099/ijsem.0.001131.

Huber, K. J., Wüst, P. K., Rohde, M., Overmann, J. & Foesel, B. U. 2014. Aridibacter famidurans gen. nov., sp. nov. and Aridibacter kavangonensis sp. nov., two novel members of subdivision 4 of the Acidobacteria isolated from semiarid savannah soil. *International Journal of Systematic and Evolutionary Microbiology*, **64**, 1866-1875, doi: 10.1099/ijs.0.060236-0.

Kembel, S. W., Cowan, P. D., Helmus, M. R., Cornwell, W. K., Morlon, H., Ackerly, D. D., Blomberg, S. P. & Webb, C. O. 2010. Picante: R tools for integrating phylogenies and ecology. *Bioinformatics*, **26**, 1463-1464, doi: 10.1093/bioinformatics/btq166.

Koboldt D. C., Zhang, Q., Larson, D. E., Shen, D., McLellan, M. D., Lin, L., Miller, C. A., Mardis, E. R., Ding, Li. & Wilson, R. K. 2012. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Research*, **22**, 568-576, doi: 10.1101/gr.129684.111.

Kumar, S., Stecher, G. & Tamura, K. 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, **33**, 1870-1874, doi: 10.1093/molbev/msw054.

Lagkouvardos, I., Joseph, D., Kapfhammer, M., Giritli, S., Horn, M., Haller, D. & Clavel, T. 2016. IMNGS: A comprehensive open resource of processed 16S rRNA microbial profiles for ecology and diversity studies. *Scientific Reports*, **6**, doi: 10.1038/srep33721.

Lane, D. 1991. 16S/23S rRNA sequencing. In: Stackebrandt, E. & Goodfellow, M. (eds) *Nucleic Acid Techniques in Bacterial Systematics.* 115-175.

Li, H. & Durbin, R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, **25**, 1754-1760, doi: 10.1093/bioinformatics/btp324.

Lozupone, C. A., Hamady, M., Kelley, S. T., Knight, R. 2007. Quantitative and qualitative beta diversity measures lead to different insights into factors that structure microbial communities. *Applied and Environmental Microbiology*, **73**, 1576-1585, doi: 10.1128/AEM.01996-06.

Lueders, T., Manefield, M. & Friedrich, M. W. 2003. Enhanced sensitivity of DNA- and rRNA-based stable isotope probing by fractionation and quantitative analysis of isopycnic centrifugation gradients. Environmental Microbiology, 6, 73-78, doi: 10.1046/j.1462-2920.2003.00536.x.

Lunau, M., Lemke, A., Walther, K., Martens-Habbena, W., & Simon, M. 2005. An improved method for counting bacteria from sediments and turbid environments by epifluorescence microscopy. *Environmental Microbiology*, 7, 961-968, doi: 10.1111/j.1462-2920.2005.00767.x.

Matthews, T. J., Borregaard, M. K., Ugland, K. I., Borges, P. A. V., Rigal, F., Cardoso, P. & Whittaker, R. J. 2014. The gambin model provides a superior fit to species abundance distributions with a single free parameter: evidence, implementation and interpretation. *Ecography*, **37**, 1002-1011, doi: 10.1111/ecog.00861.

McMurdie, P. J. & Holmes, S. 2013. Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE*, **8**, doi: 10.1371/journal.pone.0061217.

Muyzer, G., De Waal, E. C. & Uitterlinden, A. G. 1993. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Applied and Environmental Microbiology*, **59**, 695-700, doi: 0099-2240/93/030695-06$02.00/0.

Na, S. I., Kim, Y. O., Yoon, S. H., Ha, S. M., Baek, I. & Chun, J. 2018. UBCG: Up-to-date bacterial core gene set and pipeline for phylogenomic tree reconstruction. *Journal of Microbiology*, **56,** 280-285, doi: 10.1007/s12275-018-8014-6.

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., Mcglinn, D., Minchin, P. R., O'hara, R. B., Simpson, G. L., Solymos, P., Henry, M., Stevens, H., Szoecs, E. & Wagner, H. 2019. vegan: community ecology package. *R package version 2.5-6*, https://CRAN.Rproject.org/package=vegan.

Overmann, J. 2013. Chapter 7. Principles of enrichment, isolation, cultivation, and preservation of prokaryotes. *In: The Prokaryotes: Prokaryotic Biology and Symbiotic Associations*. 149–207, doi: 10.1007/978-3642-30194-0_7.

Pascual, J., Wüst, P. K., Geppert, A., Foesel, B. U., Huber, K. J. & Overmann, J. 2015a. Novel isolates double the number of chemotrophic species and allow the first description of higher taxa in

Acidobacteria subdivision 4. *Systematic and Applied Microbiology*, **38**, 534-544, doi: 10.1016/j.syapm.2015.08.001.

Pascual, J., Wüst, P. K., Geppert, A., Foesel, B. U., Huber, K. J. & Overmann, J. 2015b. *Terriglobus albidus* sp. nov., a member of the family *Acidobacteriaceae* isolated from Namibian semiarid savannah soil *International Journal of Systematic and Evolutionary Microbiology*, **65**, 3297-3304, doi: 10.1099/ijsem.0.000411.

Pruesse, E., Peplies, J. & Glöckner, F. O. 2012. Sina: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics*, **28**, 1823-1829, doi: 10.1093/bioinformatics/bts252.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J. & Glöckner, F. O. 2013. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Research*, **41**, doi: 10.1093/nar/gks1219.

R Core Team. 2019. R: a language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*, **0**, {ISBN} 3-900051-07-0, doi: http://www.R-project.org/.

Sasser, M. 2001. Identification of Bacteria by Gas Chromatography of Cellular Fatty Acids, MIDI Technical Note 101.

Tanizawa, Y., Fujisawa, T. & Nakamura, Y. 2018. DFAST: a flexible prokaryotic genome annotation pipeline for faster genome publication. *Bioinformatics*, **34**, 1037-1039, doi: 10.1093/bioinformatics/btx713.

Taylor, W. I. & Achanzar, D. 1972. Catalase test as an aid to the identification of Enterobacteriaceae. *Applied Microbiology*, **24**, 58-61.

Tindall, B. J. 1990. Lipid composition of *Halobacterium lacusprofundi*. *FEMS Microbiology Letters*, **66**, 199-202.

Tindall, B. J., Sikorski, J., Smibert, R. A. & Krieg, N. R. 2007. Chapter 15 - Phenotypic characterization and the principles of comparative systematics. *In*: Reddy, C. A., Beveridge, T. J., Breznak, J. A., Marzluf, G. & Schmidt, T. M. (editors). *Methods for General and Molecular Microbiology, 3rd ed.* American Society for Microbiology, Washington DC, 330-393, doi: 10.1128/9781555817497.ch15.

Tschech, A. & Pfennig, N. 1984. Growth yield increase linked to caffeate reduction in Acetobacterium woodii. *Archives of Microbiology*, **137**, 163-167, doi: 10.1007/BF00414460.

Turner, S., Pryer, K. M., Miao, V. P. & Palmer, J. D. 1999. Investigating deep phylogenetic relationships among cyanobacteria and plastids by small subunit rRNA sequence analysis. *The Journal of eukaryotic microbiology*, **46**, 327-338, doi: 10.1111/j.1550-7408.1999.tb04612.x.

Vieira, S., Luckner, M., Wanner, G. & Overmann, J. 2017. *Luteitalea pratensis* gen. nov., sp. nov. a new member of subdivision 6 Acidobacteria isolated from temperate grassland soil. *International Journal of Systematic and Evolutionary Microbiology*, **67**, 1408-1414, doi: 10.1099/ijsem.0.001827.

Wickham, H. 2016. *ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York*, {ISBN} 978-3-319-24277-4, doi: https://ggplot2.tidyverse.org.

Wüst, P. K., Nacke, H., Kaiser, K., Marhan, S., Sikorski, J., Kandeler, E., Daniel, R. & Overmann, J. 2016. Estimates of soil bacterial ribosome content and diversity are significantly affected by the nucleic acid extraction method employed. *Applied and Environmental Microbiology*, **82**, 2595-2607, doi: 10.1128/AEM.00019-16.

Yoon, S. H., Ha, S. M., Kwon, S., Lim, J., Kim, Y., Seo, H. & Chun, J. 2017. Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *International Journal of Systematic and Evolutionary Microbiology*, **67**, 1613-1617, doi: 10.1099/ijsem.0.001755.

Yoon, S. H., Ha, S. M., Lim, J., Kwon, S. & Chun, J. A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie van Leeuwenhoek*, **110**, 1281-1286, doi: 10.1007/s10482-017-0844-4.

Yu, G., Smith, D. K., Zhu, H., Guan, Y. & Lam, T. T-Y. ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution,* **8,** 28-36, doi: 10.1111/2041-210X.12628.

Chapter 4 – Total and active bacterial communities associated with *Hypericum* plants

4.1 – Results

4.1.1 – Scoping studies of bacterial communities associated with *Hypericum*

In 2016, an investigation of bacterial communities in bulk soil and the rhizosphere of wild *H. perforatum* and *H. polyphyllum*, and additional rhizosphere of greenhouse-cultivated *H. perforatum* was conducted (Table 1), as a step closer in understanding potential link between bacterial community and hypericin and hyperforin production in the plant genus *Hypericum.* This is referred as the first scoping study where the plants were taken from the Leibniz Institute of Plant Biochemistry (IPB) Halle. The rhizosphere bacterial community was solely investigated without additional rhizoplane or root (endophytic) bacterial community (see material and methods for details). Both total (DNA-based) and metabolically active (RNA-based) bacterial communities were investigated.

Since both of the plant species are hypericin and hyperforin-producing species, comparison between producer and non-producer to identify potential key players in the production of hypericin and hyperforin could not be conducted (Table 1). In addition, the number of biological replicates associated to bulk soil (with a maximum of two) was insufficiently low (even missing for the greenhouse-cultivated *H. perforatum*) to obtain trustworthy results from statistical tests (Table 1) when comparing the rhizosphere and bulk soil bacterial communities. Thus, the second sampling was executed in 2018 in the Botanic Garden and Botanical Museum Berlin-Dahlem (Table 2), and referred as the second scoping study.

Table 1 – Samples of the first scoping study taken in 2016 from the Leibniz IPB Halle.

Plant species	Origin	Rhizosphere samples	Sequencing platform	Biological replicates			
				Bulk soil		Rhizosphere	
				DNA	RNA	DNA	RNA
H. perforatum	Field no. 1	Rhizosphere only	HiSeq	2	2	3	3
H. perforatum	Greenhouse	Rhizosphere only	HiSeq	0	0	3	3
H. polyphyllum	Field no. 2	Rhizosphere only	HiSeq	1	1	3	3

In the second scoping study (Table 2), bulk soil and the rhizosphere bacterial communities of wild *H. perforatum* and *H. androsaemum*, and greenhouse-cultivated *H. balearicum* were inspected, each with appropriate number of biological replicates (6 biological replicates of each *H. perforatum* and *H. androsaemum* and 5 biological replicates of *H. baleaericum*). The rhizosphere bacterial community

was combined with root-endophytes and rhizoplane bacterial communities in the second scoping study (Table 2; see material and methods for details).

The result revealed that the bulk soil bacterial communities were already distinct for each plant species, which is likely due to different soil edaphic properties associated to different sampling locations. Thus, the plant species effect could not be fully determined since it was confounded with the soil effect. Therefore, a controlled-greenhouse experiment with a full-factorial crossed design was performed subsequently and will be discussed later in the next sub-chapter (see 4.1.2), involving investigation of the rhizosphere and roots bacterial communities of hypericin and hyperforin producers (*H. perforatum* and *H. olympicum*) and non-producer (*H. balearicum*) that were cultivated in two different soil substrates with distinct pH (acidic and alkaline). Both total (DNA based) and metabolically active (RNA-based) bacterial communities would be investigated and the impact of soil substrates and plant species would be determined.

Table 2 – Samples of the second scoping study taken in 2018 from the BGBM Berlin.

Plant species	Origin	Rhizosphere samples	Sequencing platform	Biological replicates			
				Bulk soil		Rhizosphere	
				DNA	RNA	DNA	RNA
H. perforatum	Field no. 1	Mixed with roots	Nextseq	6	6	6	6
H. androsaemum	Field no. 2	Mixed with roots	Nextseq	6	6	6	6
H. balearicum	Greenhouse	Mixed with roots	Nextseq	5	5	5	5

Although the comparison of bacterial community between different plant species in the second scoping study was not possible due to the reason mentioned above, the comparison between bulk soil and the rhizosphere communities for each plant species was still possible. Thus, the total (DNA-based) and metabolically active (RNA-based) bacterial communities were investigated and for each plant species, the rhizosphere bacterial community were compared to those of bulk soil.

The data of the first and second scoping studies were merged and analysed together. After processing of the raw reads, 18,616,735 reads were obtained, belonging to 72,419 sequence variants that were distributed across 92 different samples. The rarefaction curves were observed to be almost-saturated (Figure 10), which implied that our sequences cover most of the taxa present in the samples with sample coverage estimates ranging between 95.3% and 99.9%. The data was further rarified (normalized) to 71,494 reads per sample to allow fair comparison between bulk soil and rhizosphere bacterial communities.

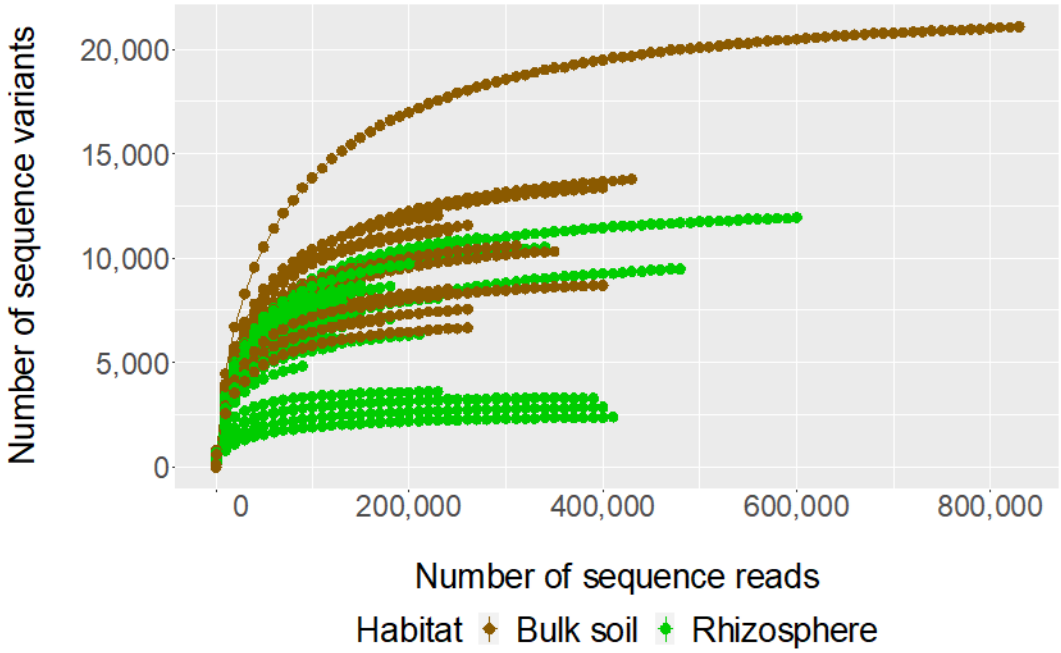


Figure 10 – Rarefaction curves for bulk soil and rhizosphere samples at sequence variant level.

As we had scoping studies (merged as one dataset) comprising wild and cultivated *Hypericum* plants and a controlled-greenhouse experiment with full-factorial crossed design, we investigated the portion of sequence variants that were shared in both datasets to have the first glimpse of how different the bacterial communities are between different soil associated with distinct locations. We found that 34,052 sequence variants (47.6% and 37.3% of sequence variants detected in the scoping studies and the greenhouse experiment, respectively) were shared between the scoping studies and the controlled-greenhouse experiment. These sequence variants that were detected in both datasets were high in abundance, contributing to 81.8% and 80.8% of the total reads in the scoping studies (5,042,498 out of 6,162,068 reads) and the controlled-greenhouse experiment (8,187,480 out of 10,136,589 reads), respectively. Since more than 50% of sequence variants in both datasets were contributing to less than 20% of total reads altogether, we could assume that many taxa belong to rare biosphere (taxa with low abundance), which is quite common in soil habitat.

Alpha diversity estimates, including richness, Shannon diversity, evenness, and alpha gambin were calculated at sequence variant level (Figure 11) to determine differences in bacterial diversity among groups of bulk soil and the rhizosphere of a particular plant species that were sampled from different locations. Of the first scoping study (samples from Halle), the total bacterial communities (DNA-based) had higher Shannon diversity compared to the active ones (RNA-based), specifically for greenhouse samples, implying a higher number of dormant taxa and/or extracellular DNA in the greenhouse soils. In addition, evenness was also higher for the total bacterial communities (DNA-based) compared to the active ones (RNA-based). This implies that taxa dominance was more

prevalent in the active bacterial communities when compared to the total ones, which was also supported by lower alpha gambin values. For alpha gambin, low values indicate a logseries-like species distribution, implying a higher level of taxa dominance whereas high values indicate a lognormal-like species distribution. The higher degree of dominance observed on the active bacterial communities when compared to the total ones in the rhizosphere may be a consequence of various selective pressures including host plant species.

Of the second scoping study (samples from Berlin), the rhizosphere of *H. perforatum* had lower Shannon diversity and evenness compared to bulk soil, specifically for the active bacterial communities (RNA-based), whereas the rhizosphere of *H. balearicum* exhibited lower alpha diversity measures (richness, Shannon diversity, and evenness) compared to the corresponding bulk soil for both total (DNA-based) and active (RNA-based) bacterial communities. The result may reflect the ability of *H. perforatum* and *H. balearicum* in selecting subset of bacterial taxa in the rhizosphere that are more favorable for them. No significant differences could be observed between bulk soil and the rhizosphere of *H. androsaemum.* Similar to what had been observed earlier, a stronger taxa dominance was observed on the active bacterial communities (RNA-based) compared to the total ones (DNA-based) in the rhizosphere of *H. perforatum* and *H. androsaemum*, as suggested by lower alpha gambin values of the active communities. This result implies that *H. androsaemum* still had an impact on its rhizosphere bacterial communities, since changes could be observed between the total (DNA-based) and active (RNA-based) bacterial communities.

Differences in bacterial community structure and composition among distinct bulk soil and rhizosphere groups (belonging to different plant species, origins, and sampling locations) were illustrated on the NMDS ordination plot based on weighted UniFrac distances (Figure 12). The NMDS stress value indicated a fair ordination in 3 dimensions that is not based upon random process. Samples were clustered mostly based on sampling locations (Halle and Berlin, which represent the first and second scoping study, respectively), with a clear separation observed between field and greenhouse samples. The results confirm that the comparison between different plant species will not be conclusive, since already the bulk soil bacterial communities were different for each plant species. Thus, the plant species effect will be confounded with the soil effect. Even the same plant species was separated based on locations, as depicted with *H. perforatum*, implying that the soil effect was stronger than the plant effect. Separation between bulk soil and the rhizosphere bacterial communities was less pronounced, suggesting that the impact of host plant in driving the rhizosphere communities was weaker than the soil effect that corresponds to different sampling locations.

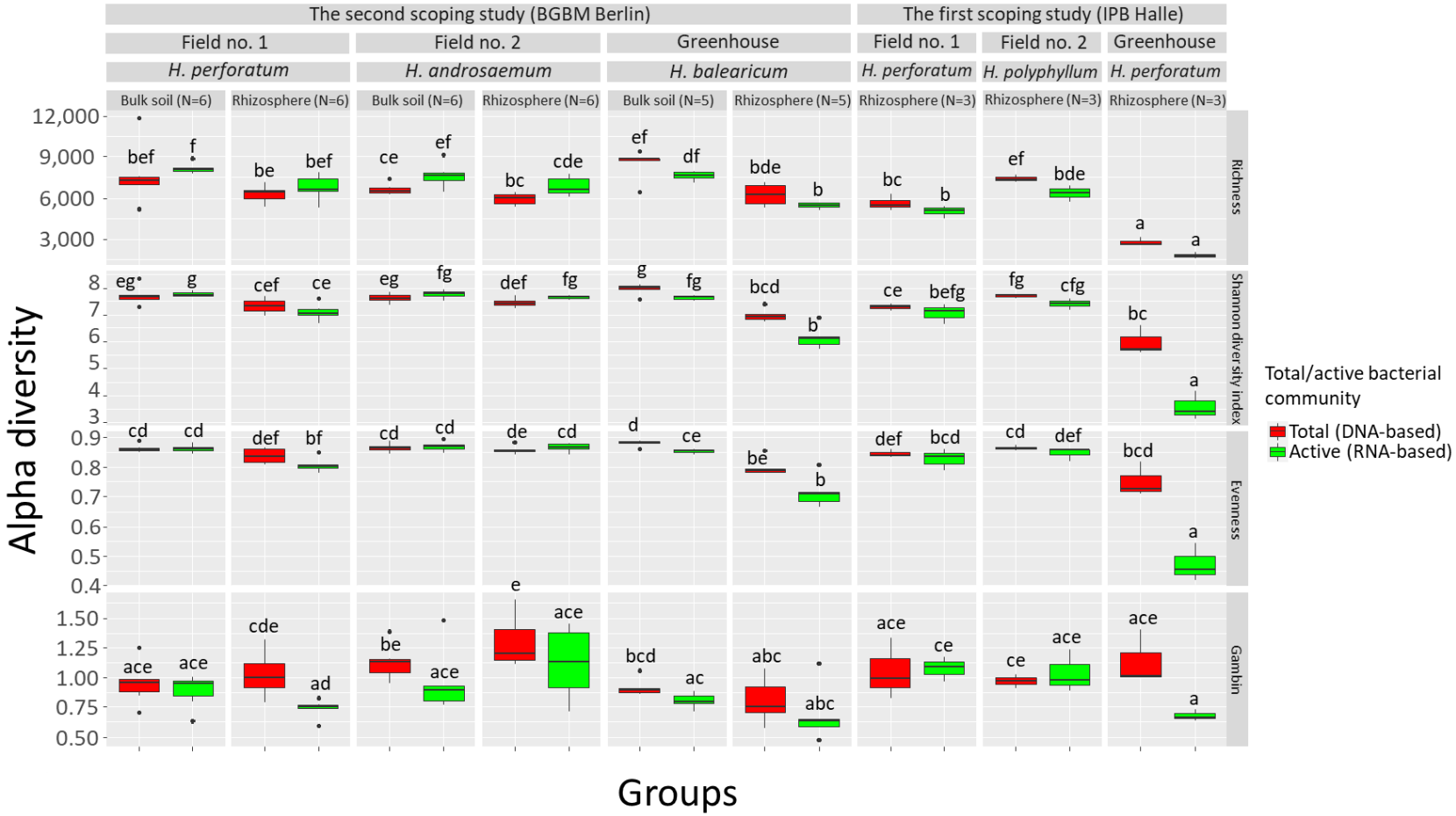


Figure 11 – Alpha diversity measures for bacterial communities in the scoping studies at sequence variant level. Both total (DNA-based; red boxplots) and active (RNA-based; green boxplots) bacterial communities were investigated. Significant differences between distinct habitat types (bulk soil and the rhizosphere), plant species (*H. perforatum*, *H. androsaemum, H. balearicum,* and *H. polyphyllum*), origins (fields and greenhouses), and sampling locations (Berlin for the second scoping study and Halle for the first scoping study) were denoted as letters at the top of each box plot ($p < 0.05$, multcomp test). N represents the number of samples for each group.

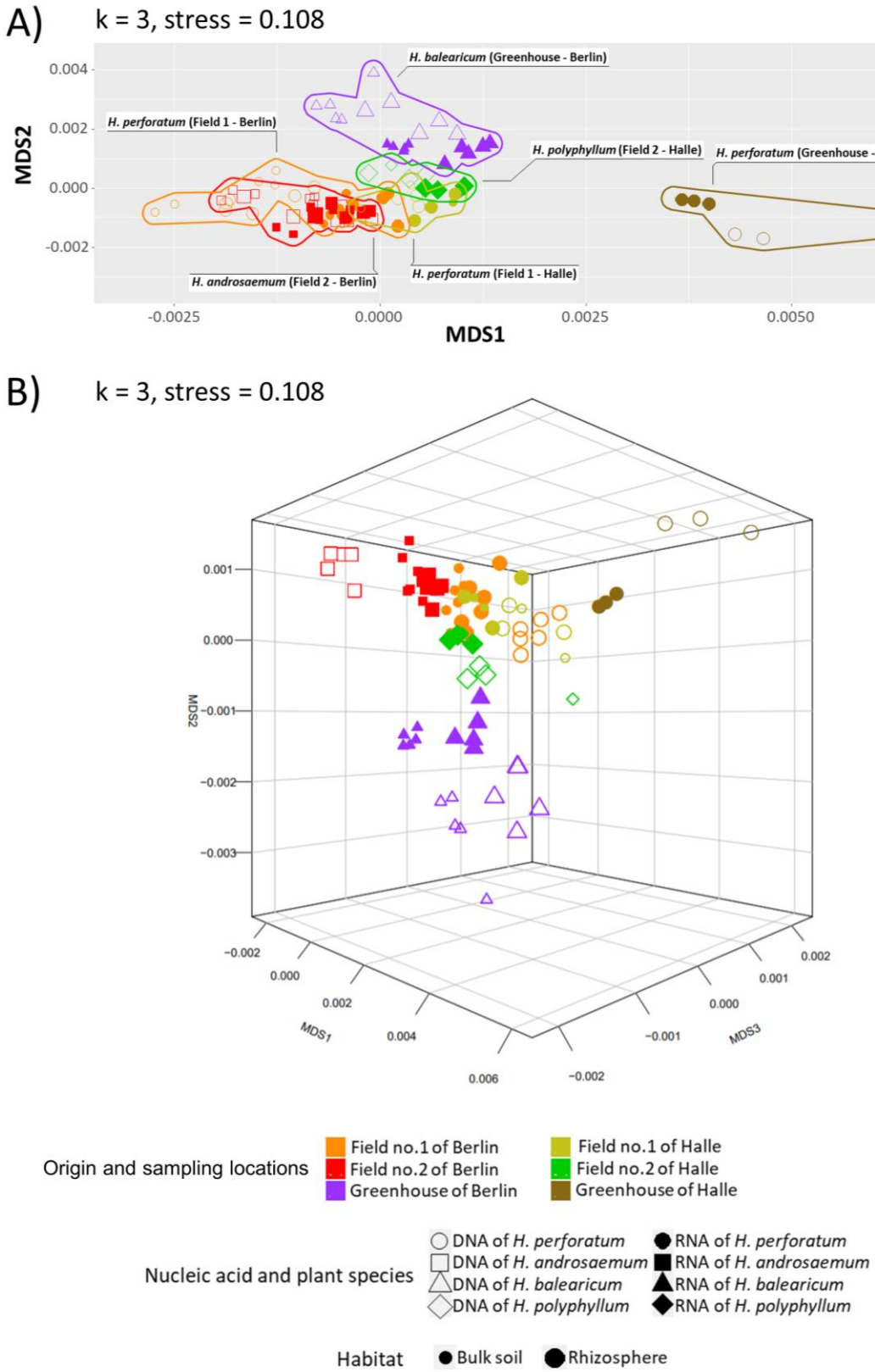


Figure 12 – Non-metric multidimensional scaling (NMDS) plots based on weighted UniFrac distances along the first and second axis (A) and with the addition of third axis (B), illustrating separation of samples according to the structure and composition of the bacterial communities.

Moreover, Halle-greenhouse samples (from the first scoping study) were quite distinct from the rest, possibly due to the heat sterilization process applied to the soil in prior, as also supported by the lowest alpha diversity measures. In general, separation between the total (DNA-based) and active (RNA-based) bacterial communities was evident regardless of sampling locations (Halle/Berlin), origins (fields/greenhouses), plant species (*H. perforatum*, *H. androsaemum, H. balearicum,* and *H. polyphyllum*), and habitat types (bulk soil/the rhizosphere), implying different level of transcriptional activities among the bacteria.

Bacterial community composition was further investigated at phylum level to determine differences between different habitat types, plant species, origins and sampling locations (Figure 13). The result revealed that greenhouse samples of Halle (the first scoping study), that were separated quite far away from the rest of the samples on the ordination plot*,* contain the highest Cyanobacterial reads when compared to other groups. Most of the Cyanobacterial reads were represented by 4 unique sequence variants belonging to the family *Nostocaceae*, which were responsible for 98.6 ± 0.7% and 99.7 ± 0.1% of Cyanobacterial reads in the total (DNA-based) and active (RNA-based) bacterial communities, respectively (Figure 14). This implies that specific soil edaphic properties of Halle-greenhouse samples may select for those 4 dominant sequence variants*.* Of field samples of Halle (the first scoping study), the taxa abundance at phylum level seemed to be similar between distinct plant species and habitat types (Figure 13). Since each plant species harbour distinct bacterial communities based on separation on the ordination plot, those differences may be observed at deeper taxonomic levels. Greenhouse soils of Halle seemed to favor Actinobacteriota, as suggested with higher relative abundance compared to other soil groups, but this is not the case for Proteobacteria and Acidobacteriota.

Of samples from the second scoping study (collected from Berlin), bulk soil of greenhouse appeared to accommodate Chloroflexi and Firmicutes when compared to other bulk soil and the rhizosphere groups, and hindered Proteobacteria when compared to the respective rhizosphere samples. Of field samples of the second scoping study (collected from Berlin), bulk soil of *H. perforatum* also hindered Proteobacteria when compared to the respective rhizosphere samples but this is not the case for *H. androsaemum.* Moreover, the abundance of Planctomycetota was higher in the active communities (RNA-based) of *H. androsaemum* when compared to the total ones (DNA-based), implying high transcriptional activity of the members. However, this trait may be more related to soil edaphic properties of field no.2 of the second scoping study (where *H. androsaemum* samples were collected) instead of plant effect, since the pattern was observed for both bulk soil and rhizosphere bacterial communities.

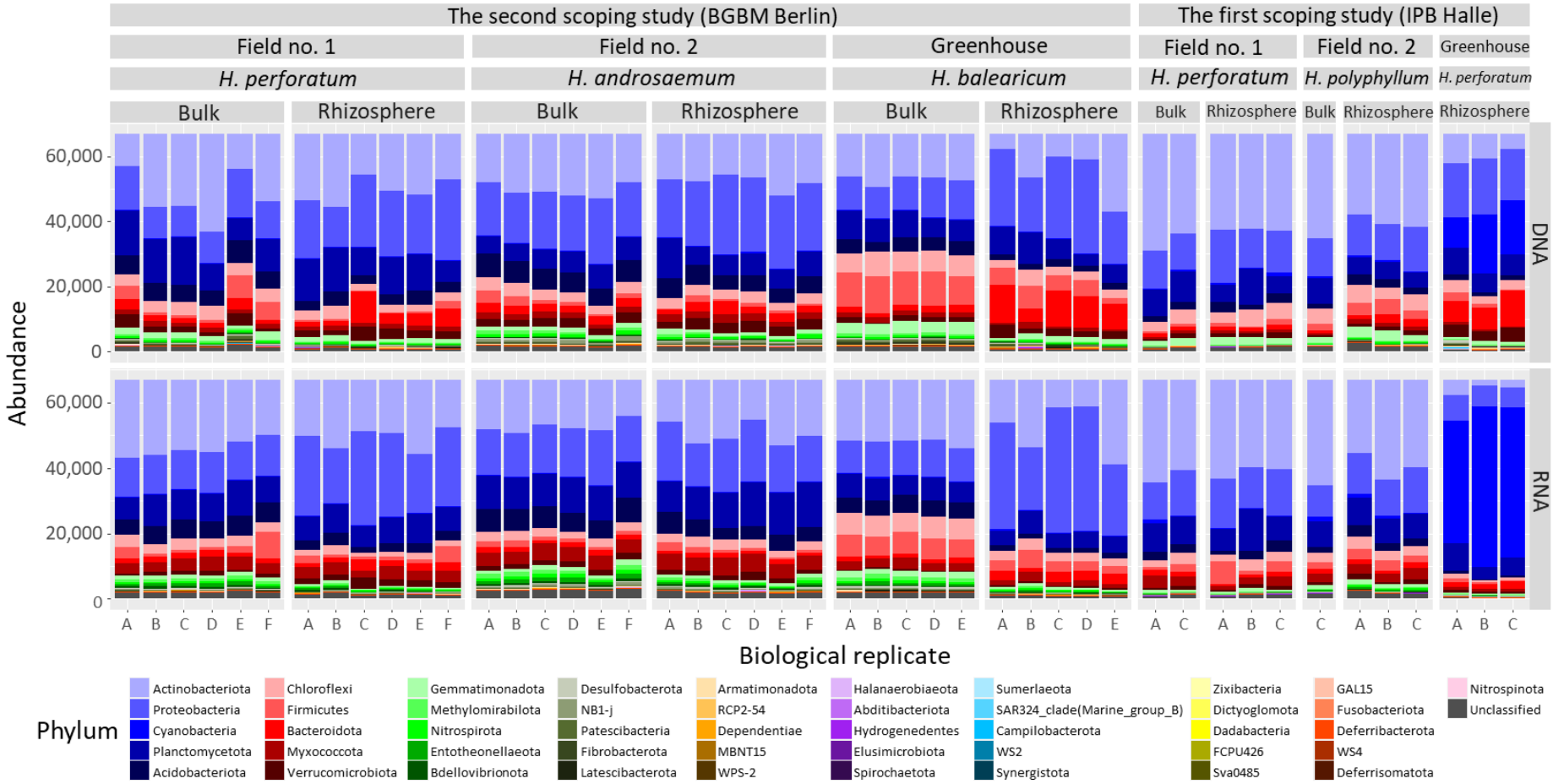


Figure 13 – Taxonomic composition of bacterial communities at phylum level in bulk soil and the rhizosphere of *H. perforatum, H. androsaemum, H. balearicum* and *H. polyphyllum* from the first (collected from IPB Halle) and second (collected from BGBM Berlin) scoping studies.

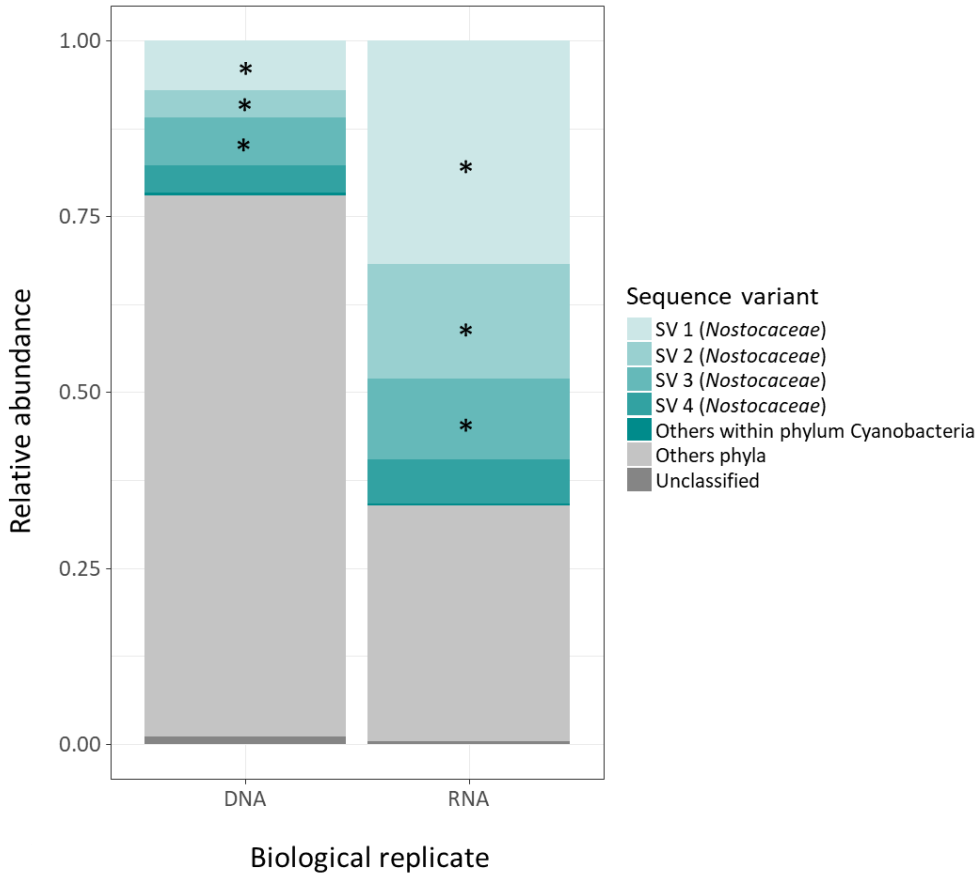


Figure 14 – Bacterial community composition specifically for greenhouse samples of the first scoping study (collected from IPB Halle), with a special focus on the 4 dominant sequence variants belonging to the family *Nostocaceae.* Significantly enriched sequence variants between the total (DNA-based) and active (RNA-based) bacterial communities were determined via multcomp test ($p < 0.01$) and were masked with asterisk.

In general, Firmicutes was favoured in bulk soil of Berlin greenhouse (correspond to bulk soil of *H. balearicum*) when compared to other bulk soil and the rhizosphere groups, regardless of sampling locations (Halle or Berlin), and origins (fields or greenhouses). Moreover, the rhizosphere of *H. balearicum* and *H. perforatum* collected from Berlin and Halle greenhouses, respectively, promoted Bacteroidota when compared to other bulk soil and rhizosphere groups, regardless of the sampling locations (Halle or Berlin). However, this was only observed for the total communities and thus may represent dormant members of Bacteroidota or relic DNA. The abundance of Myxococcota was generally higher in the active communities (RNA-based) when compared to the total ones (DNA-based) except for the greenhouse samples of Halle, implying that the members may play important roles related to soil ecological funtions as they were active across distinct soil samples. To conclude, we observed differences in bacterial community composition even at phylum level between bulk soil and the rhizosphere groups belonging to distinct plant species (*H. perforatum, H. androsaemum, H.*

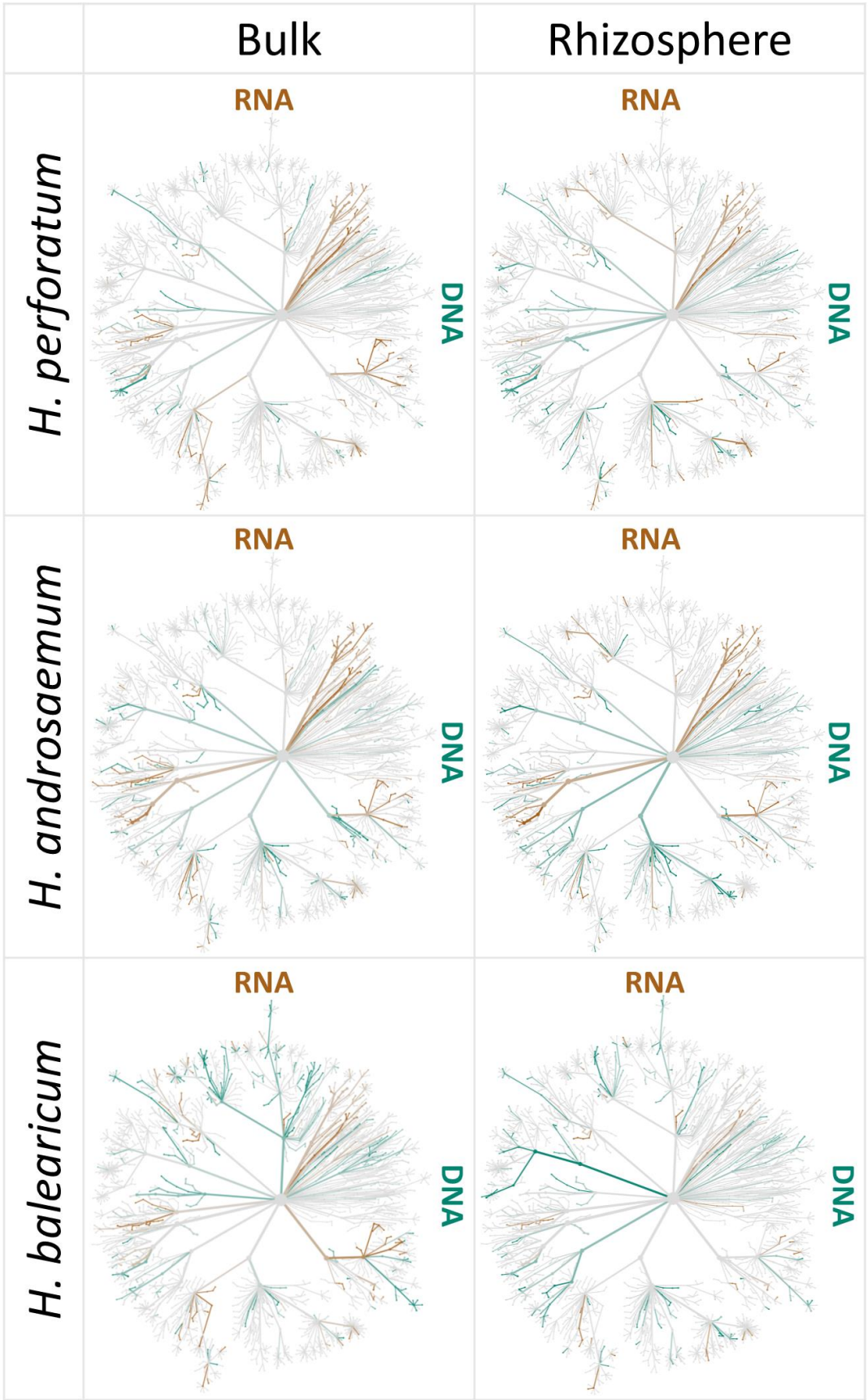
balearicum, and *H. polyphyllum*), origin (fields and greenhouses), and sampling locations (Halle and Berlin). However, we could not determine the degree of impact of different potential drivers such as soil and plant species, as the plant species effect is likely confounded with the soil effect since the bulk soil bacterial communities already different for each plant species.

In order to identify bacterial taxa that were enriched in the active bacterial communities (RNA-based) when compared to the total ones (DNA-based), suggesting higher transcriptional activities that are likely relevant to the soil or plant ecological functions, comparison between the active (RNA-based) and total (DNA-based) bacterial communities was conducted for each bulk soil and rhizosphere samples of a particular plant species (Figure 15). Since some bulk soil samples were missing or not available at all, samples from the first scoping study (all samples collected from Halle) were excluded from the analysis from here onwards, unless mentioned otherwise. In addition, samples from the second scoping study (collected from Berlin; Table 2) have higher number of biological replicates compared to those of the first scoping study (collected from Halle; Table 1) and thus were also selected for the analysis on the basis of a higher statistical power.

The enriched taxa in the active bacterial communities (RNA-based) compared to the total ones (DNA-based), independently of the habitat types (bulk soil and rhizosphere) and plant species (*H. perforatum, H. androsaemum,* and *H. balearicum*), included OLB14 of Chloroflexi, *Solirubrobacter*, *Haliangium*, "*Candidatus* Solibacter", "*Candidatus* Entotheonella", OM27 clade of *Bdellovibrionaceae, Pajaroellobacter*, *Anaeromyxobacter*, *Bryobacter*, and *Polyangium* (Figure 15). These taxa may be important players in soil ecosystem functioning as they were enriched in the active communities (RNA-based) of diverse soils and not necessarily linked to a specific plant species. On the other hand, CCD24 and PLTA13 of *Gammaproteobacteria, Candidatus* Magasanikbacteria, *Candidatus Peribacteria,* LWQ8 of *Saccharimonadales,* Pir4 lineage of *Pirellulaceae, Dongia*, MND1 and Ellin6067 of *Nitrosomonadaceae, Luteolibacter*, *Caldalkalibacillus*, *Parcubacteria*, and *Tumebacillus* were among the enriched in the total (DNA-based) bacterial communities (Figure 15). Their enrichment in the total bacterial communities (DNA-based) compared to the active ones (RNA-based) suggests that they represent dormant taxa or extracellular DNA.

TK10 of Chloroflexi, *Paenibacillus* and *Anaerocolumna* were consistently enriched in the active bacterial communities (RNA-based) compared to the total ones (DNA-based), regardless of plant species and only in the rhizosphere. This implies that the taxa are important for the host plant in general but not necessarily related to specific ecological functions that are performed by a particular plant species.

A)


Bulk
Rhizosphere
H. perforatum
H. androsaemum
H. balearicum
RNA
DNA

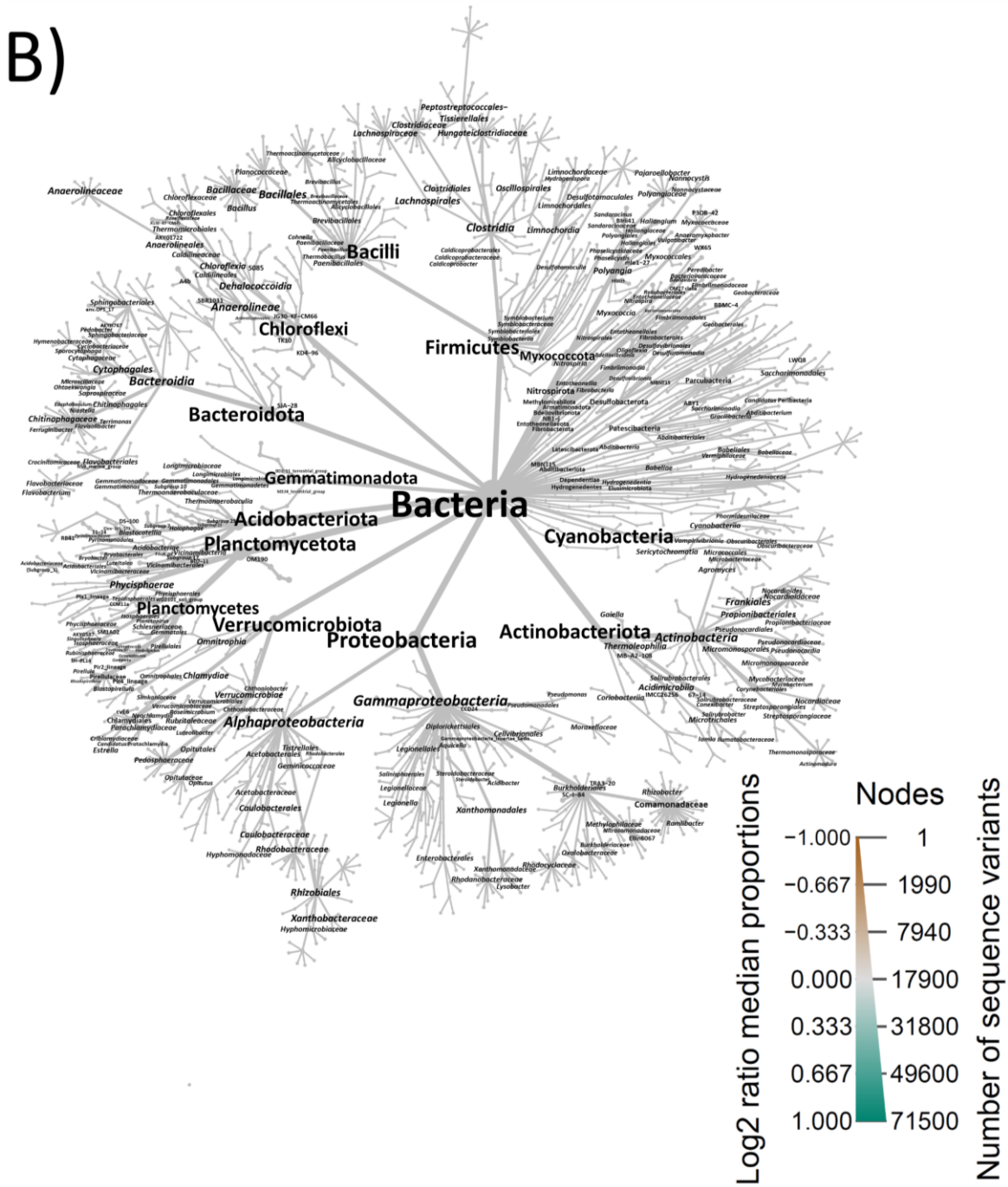


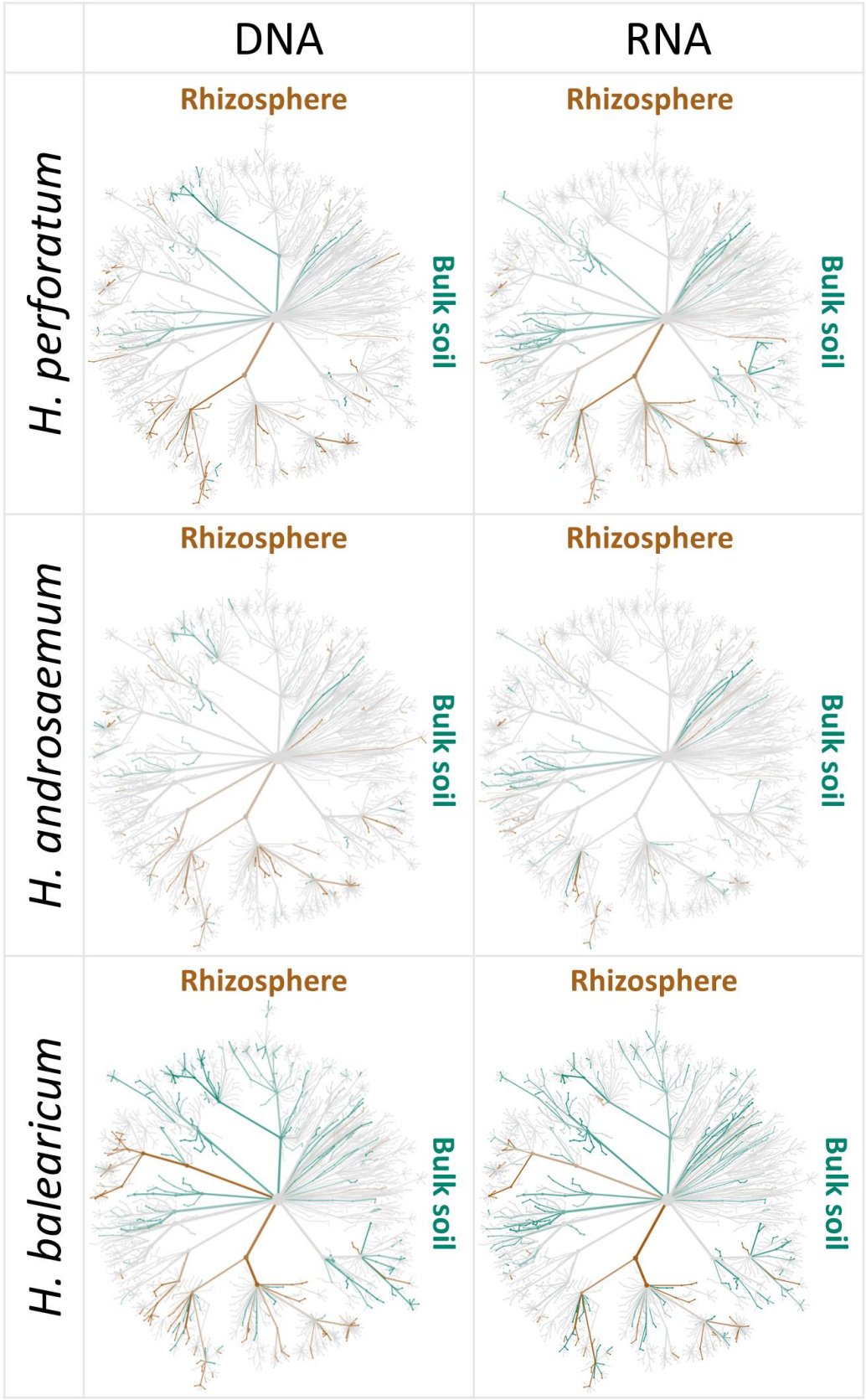
Figure 15 – Pairwise comparison of bacterial communities between the total (DNA-based) and active (RNA-based) communities of bulk soil and the rhizosphere of *H. perforatum, H. androsaemum* and *H. balearicum* from the second scoping study (collected from the BGBM Berlin; Table 2), from phylum up to genus level (A). Wilcoxon rank-sum test was performed to determine significant differences ($p < 0.05$). Significantly enriched taxa were colored according to the community type (either total or active bacterial communities) where they were found to be enriched in (A) and the color intensity reflects log-2 ratio in median proportions (the values are depicted in Figure B). Figure B depicts the key tree with the taxonomic information. Node diameter reflects number of sequence variant classified as that taxon whereas edge width reflects number of reads.

Among all taxa that were enriched in the active bacterial communities (RNA-based) compared to the total bacterial communities (DNA-based); some were specifically enriched in the rhizosphere of a specific plant species. These included PMMR1 of *Caulobacteraceae, Microlunatus, Aeromicrobium, Conexibacter, Bosea, Actibacterium, Elioraea, Oscillochloris, Pseudobacteroides, Candidimonas, Solibacillus,* and *Deferrisoma* that were specific for *H. balearicum*, BD7-11 of Planctomycetota, IMCC26256 of *Acidimicrobiia,* Eel-36e1D6 of Polyangiales, *Agromyces, Marmoricola, Phenylobacterium, Microbacterium, Zavarzinella, Gemmata,* IMCC26207 of *Microtrichaceae, Gemmatimonas, Ureibacillus, Sporosarcina, Tagaea, Nitrosospira, Pseudenhygromyxa, Sericytochromatia, Minicystis,* ADurb.Bin118 of *Pedosphaeraceae*, "*Candidatus* Anammoximicrobium", *Cephaloticoccus,* "*Candidatus* Obscuribacter", "*Candidatus* Nostocoida", and *Myxococcus* that were specific for *H. androsaemum,* and JG30-KF-CM45 of *Thermomicrobiales,* Blfdi19 of *Polyangia, Pseudomonas*, *Mesorhizobium*, *Microvirga*, wb1-P19 of *Nitrosococcaceae, Actinomycetaspora, Rhizobacter, Variovorax, Phaselicystis, Friedmanniella,* Tychonema CCAP 1459-11B of *Phormidiaceae, Luteimonas, Adhaeribacter, Phaeodactylibacter, Edaphobaculum, Catelliglobosispora, Ideonella, Sporocytophaga, Archangium, Syntrophorhabdus, Peredibacter, Thermoflavimicrobium,* and *Euzebya* that were specific for *H. perforatum*. Since these bacterial taxa were not enriched in the active communities (RNA-based) of the corresponding bulk soil, they may perform specific functions that are relevant to the host plant.

Furthermore, comparison of bacterial communities between bulk soil and the rhizosphere of *H. perforatum, H. androsaemum,* and *H. balearicum* from the second scoping study (collected from Berlin; Table 2) was conducted to determine rhizosphere-enriched taxa that are likely important for each plant species (Figure 16). Taxa enrichment was quite evident between bulk soil and the rhizosphere of *H. balearicum* compared to the other plant species, which was implied by stronger color differences on the plot, mainly due to higher number of bulk soil-enriched taxa when compared to those of the rhizosphere. Thus, the result not necessarily implies the selection of higher number of taxa by *H. balearicum.*

For both total (DNA-based) and active (RNA-based) bacterial communities, PMMR1 of *Caulobacteraceae, Hydrogenophaga, Dyadobacter, Ohtaekwangia, Niastella,* and *Pseudoflavitalea* were among the rhizosphere-enriched taxa regardless of plant species and thus may be important for the host plant in general. JG30-KF-CM66 of Chloroflexi, subgroup 7 of Acidobacteriota, BD2-11 terrestrial group of Gemmatimonadota, and BSV26 of *Kryptoniales* were among the bulk soil-enriched taxa regardless of the plant species, implying that these taxa do not perform functions that are crucial for the plants.

A)


DNA
RNA
H. perforatum
H. androsaemum
H. balearicum
Rhizosphere
Bulk soil

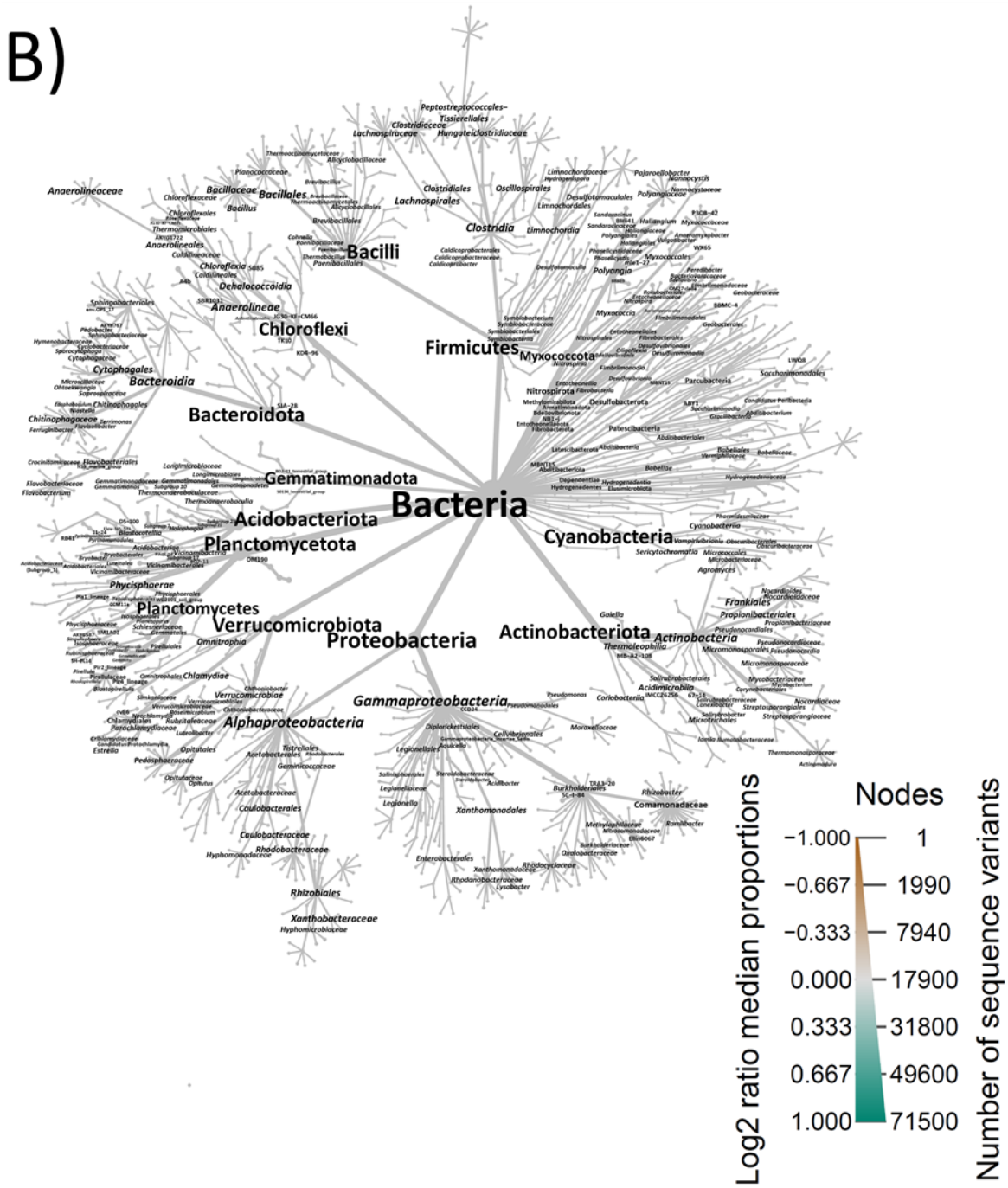


Figure 16 – Pairwise comparison of bacterial communities between bulk soil and the rhizosphere of *H. perforatum, H. androsaemum* and *H. balearicum* from the second scoping study (collected from the BGBM Berlin; Table 2) for both total (DNA-based) and active (RNA-based) communities, from phylum up to genus level (A). Wilcoxon rank-sum test was performed to determine significant differences ($p < 0.05$). Significantly enriched taxa were colored according to the habitat type (either bulk soil or the rhizosphere) where they were found to be enriched in (A) and the color intensity reflects log-2 ratio in median proportions (the values are depicted in Figure B). Figure B depicts the key tree with the taxonomic information. Node diameter reflects number of sequence variant classified as that taxon whereas edge width reflects number of reads.

The rhizosphere-enriched taxa in both total (DNA-based) and active (RNA-based) bacterial communities specifically for *H. perforatum* included *Hyphomicrobium, Pseudorhodoplanes, Sphingobium, Rhodoferax,* "*Candidatus* Xiphinematobacter", and *Pedosphaera.* Morever, Pir 4 lineage of *Pirellulaceae, Chthoniobacter, Povalibacter*, P3OB-42 of *Myxococcaceae,* and *Actinoplanes* were the rhizosphere-enriched taxa in both total (DNA-based) and active (RNA-based) bacterial communities of *H. androsaemum,* whereas *Mariniflexile, Pseudolabrys, Xenophilus, Blastopirellula, Arenibacter, Paracoccus, Taibaiella,* UBA6140 of *Methylophilaceae, Acidovorax, Paraglaciecola, Algoriphagus, Peredibacter, Hylemonella,* and *Pelagibius* were the rhizosphere-enriched taxa in both total (DNA-based) and active (RNA-based) bacterial communities of *H. balearicum.* These taxa may be crucial in maintaining host plant health and productivity and perform specific functions that are relevant to each host plant species.

The rhizosphere-enriched taxa regardless of the plant species and specifically for active bacterial communities (RNA-based) included PMMR1 of *Caulobacteraceae, Hydrogenophaga, Bosea, Dyadobacter, Ohtaekwangia, Chitinophaga, Niastella, Pseudoflavitalea,* and *Prosthecobacter.* The enrichment of these taxa in the rhizosphere only in the active (RNA-based) bacterial communities suggests that these taxa are likely active and interact with the host plant to promote their growth (unspecifically, since they were identified as rhizosphere-enriched taxa for all three plant species). On the other hand, the bulk soil-enriched taxa regardless of plant species included subgroup 2, 7, 18 and 22 of Acidobacteriota, TK10 and JG30-KF-CM66 of Chloroflexi, OM190 of Planctomycetota, BD2-11 terrestrial group of Gemmatimonadota, CCD24 of *Gammaproteobacteria,* 11-24 of *Blastocatellia,* FS118-23B-02 of *Coriobacteriia*, TRA3-20 and SC-I-84 of *Burkholderiales,* BSV26 of *Kryptoniales,* AKYH767 of *Sphingobacteriales,* MND1 and mle1-7 of *Nitrosomonadaceae, Nitrospira, Azospirillum,* RB41 of *Pyrinomonadaceae, Tagaea,* and Pir2 lineage of *Pirellulaceae.* Since they were identified as bulk-soil enriched taxa specifically in the active communities (RNA-based), they may actively contribute to soil ecosystem functions that are not related to the plant well-being.

The rhizosphere-enriched taxa specifically in the active bacterial communities (RNA-based) and only in *H. perforatum* included *Variovorax, Chryseolina,* IMCC26134 of *Opitutaceae, Polaromonas, Kineosporia, Lacunisphaera, Mucilaginibacter Legionella, Rubrivivax,* "*Candidatus* Rubidus", *Pantoea,* and DEV007 of *Verrucomicrobiales*, implying that these taxa may perform specific functions related to the ecological roles of *H. perforatum.* In addition, SH-PL14 of *Rubinisphaeraceae, Phenylobacterium, Zavarzinella,* SWB02 of *Hyphomonadaceae, Bythopirellula,* A0839 of *Rhizobiales,* CL500-29 marine group of *Ilumatobacteraceae*, *Vicinamibacter, Fimbriimonadales, Ideonella, Kouleothrix,* and *Oscillochloris* were specific for *H. androsaemum* and *Paenibacillus, Neochlamydia,*

Rhodomicrobium, MM2 of *Methylophilaceae,* and *Leptospira* were specific for *H. balearicum*, suggesting that these taxa carry out specific ecological functions that are relevant for each of the host plant species.

The rhizosphere-enriched taxa specifically in the total (DNA-based) bacterial communities regardless of the plant species included *Hydrogenophaga, Opitutus, Rhizobacter, Piscinibacter, Methylotenera, Dyadobacter, Diplosphaera, Thalassospira, Ohtaekwangia, Steroidobacter, Niastella, Pseudoflavitalea, Chryseolina,* IMCC26134 of *Opitutaceae, Verrucomicrobium, Polaromonas,* and *Lacunisphaera.* Since these taxa were not among the rhizosphere-enriched taxa in the active communities (RNA-based), they may represent dormant taxa or relic DNA. The bulk soil-enriched taxa specifically in the total (DNA-based) communities, regardless of the plant species, included S0134 and BD2-11 terrestrial groups of Gemmatimonadota, JG30-KF-CM66 of Chloroflexi, subgroup 7 of Acidobacteriota, IMCC26256 of *Acidimicrobiia,* BSV26 of *Kryptoniales, Bacillus, Paenisporosarcina, Ureibacillus, Brevibacillus, Thermoactinomyces, Aneurinibacillus,* and *Tuberibacillus.* These bacterial taxa are most likely dormant or representing relic DNA in the soil samples.

Moreover, *H. perforatum*-specific rhizosphere-enriched taxa specifically in the total (DNA-based) bacterial communities included *Reyranella, Phenylobacterium, Haliangium, Nordella,* "*Candidatus* Chloroploca", *Polyangium, Aridibacter, Haloferula, Georgfuchsia, Oscillochloris, Candidatus* Vogelbacteria, and DEV008 of *Pedosphaeraceae.* In addition, SAR324 clade, Ga0077536, WD260, and R7C24 of *Gammaproteobacteria,* env.OPS_17 of *Sphingobacteriales,* LWQ8 of *Saccharimonadales, Dongia,* MND1 of *Nitrosomonadaceae, Acidibacter, Falsirhodobacter, Mucilaginibacter, Stella, Candidatus* Nomurabacteria, *Candidatus* Adlerbacteria, *Alkalispirochaeta*, and *Litorilinea* were specific for *H. androsaemum* and *Pseudomonas, Pedomicrobium, Flavobacterium, Bosea, Dokdonella, Curvibacter, Aureimonas, Thermomonas, Achromobacter, Sporocytophaga, Roseivirga, Pedobacter, Subsaxibacter, Thalassobaculum, Marinoscillum, Cephaloticoccus, Candidatus* Lloydbacteria, *Colwellia,* C2U of *Rhizobiales, Ereboglobus, Pelagicoccus, Reichenbachiella, Imperialbacter, Methylobacillus,* and *Ferrovibrio* were specific for *H. balearicum.* Since they were only enriched in the rhizosphere in the total (DNA-based) bacterial communities, they might only represent dormant taxa or extracellular DNA in the rhizosphere of the corresponding plant species.

In order to determine if there are any differences in the transcriptional activity among bacterial taxa at sequence variant level, the DNA abundance were plotted against RNA abundance for each individual sequence variant that were detected in the samples from the second scoping study (collected from BGBM Berlin; Figure 17A). We hypothesized that some bacterial taxa may be more

transcriptionally active (higher RNA abundance compared to DNA) and thus crucial for soil or plant ecosystem functioning. Although the result suggests that DNA and RNA abundances were moderately correlated with Spearman correlation value of 0.53, many taxa were observed with contrasting values of RNA and DNA abundances, implying different level of activities across soil or rhizospheric bacteria.

However, the contrast values of RNA and DNA abundances among bacterial taxa might occur due to random processes according to the analysis done by Johannes Sikorski of DSMZ (see sub-chapter 3.1.7.3), leading to an establishment of a specific threshold to identify the active taxa based on rRNA:rDNA ratio (blue line, Figure 17B). Taxa can be identified as active given the rRNA:rDNA ratio values above the threshold. The threshold identifies 5193 active sequence variants in the dataset comprising samples from the second scoping study (collected from BGBM Berlin; Table 2). Samples of the first scoping study (collected from Halle) were not included in the analysis due to the insufficient number of bulk soil biological replicates, which are needed subsequently to validate active taxa in a specific group.

Approximately 70% active sequence variants were detected in both total (DNA-based) and active (RNA-based) bacterial communities while the remainder were only detected in the active communities. In general, approximately 1% sequence variants were considered as active in bulk soil and the rhizosphere, regardless of plant species (*H. perforatum, H. androsaemum,* and *H. balearicum*) and origins (fields and greenhouse) (Figure 18).

In order to test the ability of the threshold to identify the same active taxa across biological replicates, the distribution of active taxa across biological replicates was compared to a stochastic null model obtained from random sampling (see sub-chapter 3.1.7.4 for detail). Random sampling of the rhizosphere samples of *H. perforatum* (included in the second scoping study that were collected from BGBM Berlin) revealed that the majority of the randomly-sampled sequence variants (~98%) only appeared in 1 biological replicate (Figure 19, boxplot on the top/upper-left corner). On the other hand, approximately 58.4% of the active sequence variants were unique for 1 biological replicate, while the rest are shared in at least 2 or even up to 6 biological replicates (Figure 19, green stacked bar on the left). The result implies that the threshold is quite reliable to identify active taxa and the taxa that were active only in 1 biological replicate are most likely selected through a random process, since the majority of randomly-sampled sequence variants belonged to only 1 biological replicate.

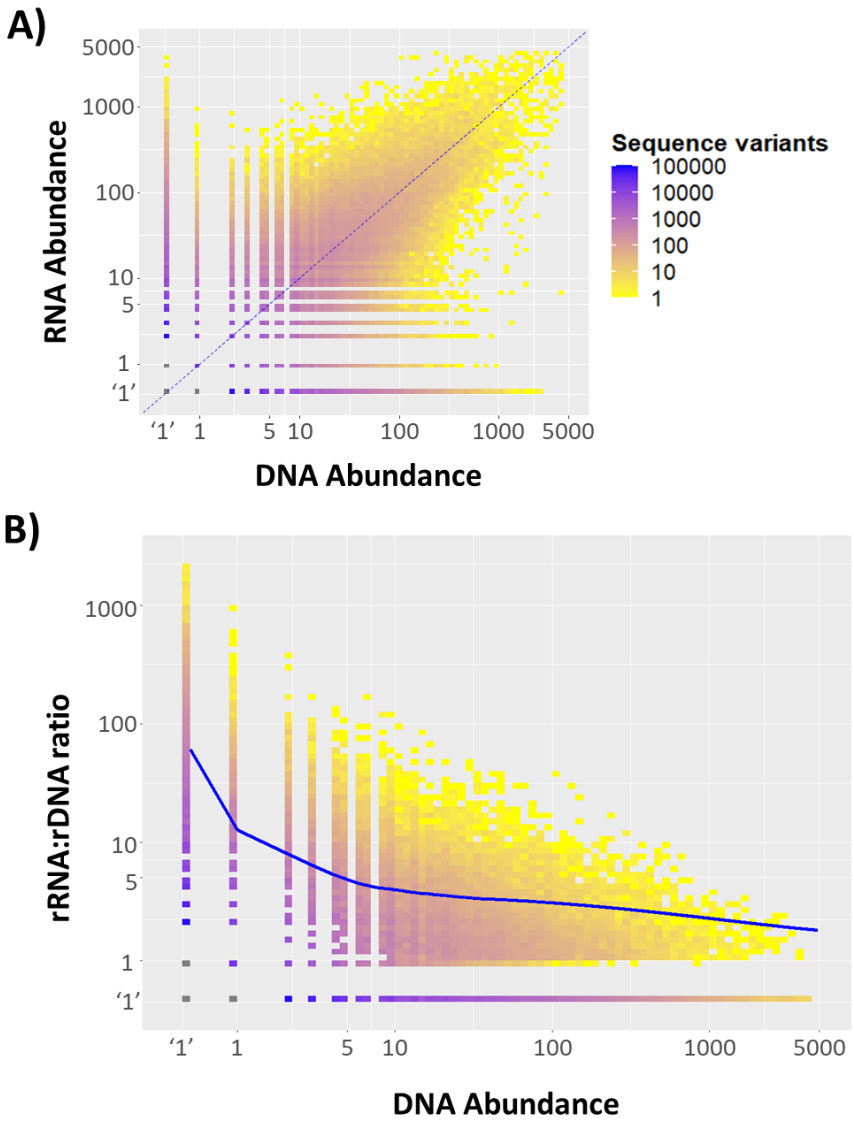


Figure 17 – DNA vs RNA abundance plot (A) for all sequence variants that were detected in the samples from the second scoping study (collected from the BGBM Berlin; Table 2). Figure B was constructed with the same data but with the rRNA:rDNA ratio values displayed as y axis. The blue line depicts a threshold to identify active taxa (active: rRNA:rDNA ratio above the threshold). A pseudocount of 1 (replacing 0 and labeled as '1') was introduced for the taxa that only present either in RNA or DNA data.

There were approximately 1.7% of randomly-sampled sequence variants that were shared in 2 biological replicates. Of the active taxa, around 15.4% were detected to be active in 2 biological replicates. Based on this, it is less likely that the active taxa observed in 2 biological replicates were selected through a random process. Therefore, detection of activity at least in 2 biological replicates was selected as a criterium to determine group-active taxa (in bulk soil or rhizosphere of a specific plant species). In addition, further increase in the number of biological replicates gave lower percentage of randomly-sampled sequence variants, suggesting that random effect is reduced as the

detection of active sequence variants in multiple biological replicates increased. The same pattern was observed in all bulk soil and rhizosphere groups, despite not shown here in the thesis.

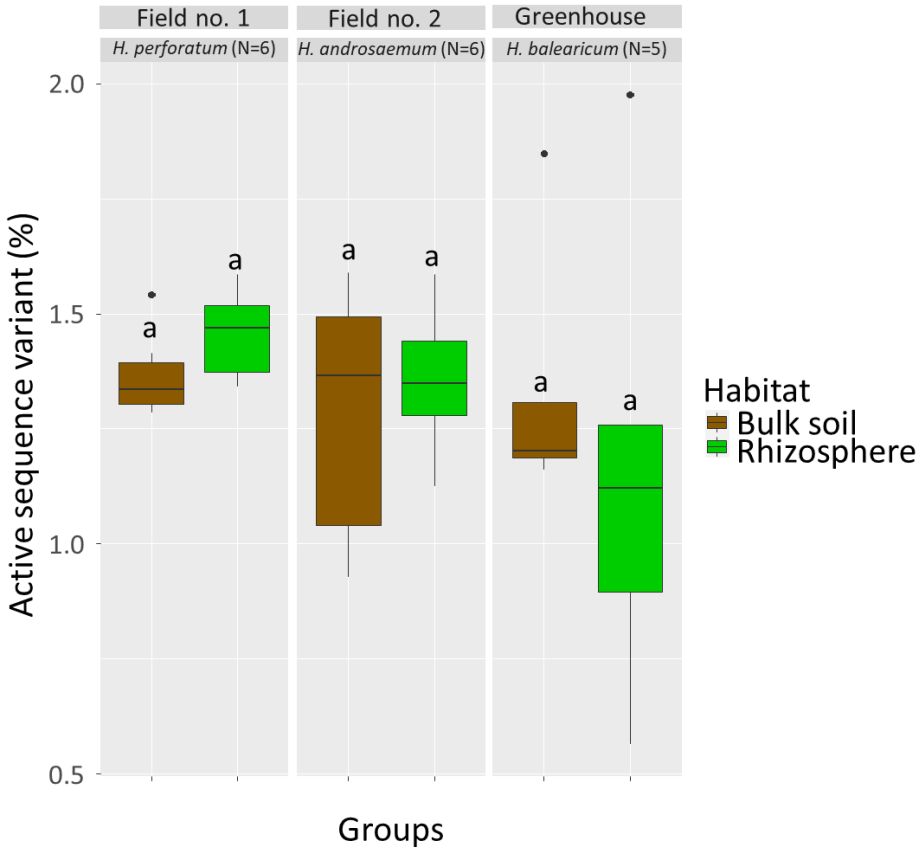


Figure 18 – Percentage of active sequence variants in bulk soil and the rhizosphere of *H. perforatum, H. androsaemum,* and *H. balearicum* from the second scoping study (collected from the BGBM Berlin; Table 2). Significant differences between different habitat (bulk soil or rhizosphere) and plant species were determined by multcomp and are denoted by different letters ($p < 0.05$). N represents the number of samples for each group.

There were 657 sequence variants among the group-active taxa. Of these, approximately 46% could be identified at genus level (Figure 20). These included diverse genera belonging to 17 different phyla according to the SILVA138 database. Of *H. perforatum* samples, *Microvirga, Hyphomicrobium, Pseudomonas, Pseudonocardia, Nitrospira, Pseudoxanthomonas, Rhizobacter,* and *Niastella* were among the genera that were associated with the active sequence variants. Among those mentioned genera, *Pseudomonas, Pseudoxanthomonas, Rhizobacter* and *Niastella* were only associated to the rhizosphere while the rest were also active in bulk soil. Of *H. androsaemum* samples, the active sequence variants in the rhizosphere included taxa that were classified as *Haliangium, Nitrospira, Microlunatus, Pedomicrobium, Phenylobacterium, Inquilinus, Bacillus, Iamia,* and *Pseudomonas.* Of these, the last three were only associated to the rhizosphere while the rest were also active in bulk

soil. Lastly, *Bacillus, Aureimonas, Bosea, Hydrogenophaga, Solirubrobacter, Nocardioides, Paenibacillus,* and *Streptomyces* were among genera that were associated to the active sequence variants of *H. balearicum.* Of these, *Aureimonas*, *Bosea*, *Hydrogenophaga,* and *Paenibacillus* were active only in the rhizosphere. For each plant species, the sequence variants that are specifically active in the rhizosphere may be key players in maintaining plant health and productivity or perform specific ecological functions related to the plant species.

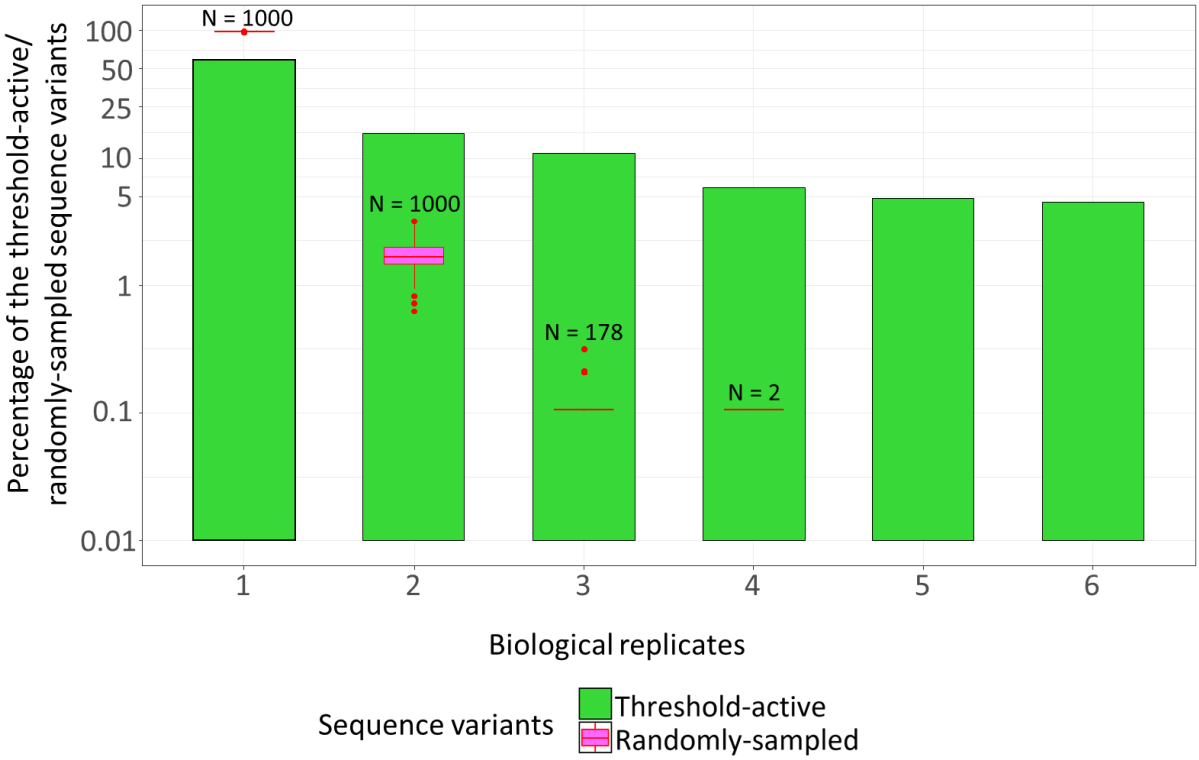
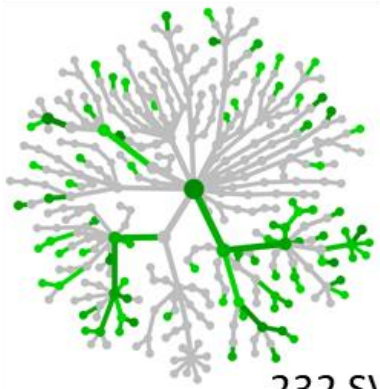
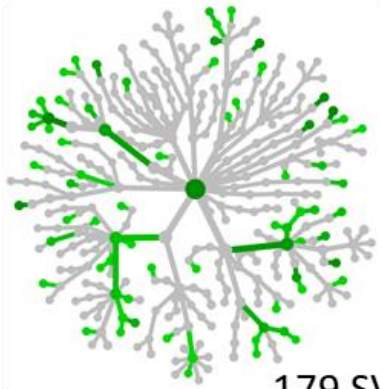
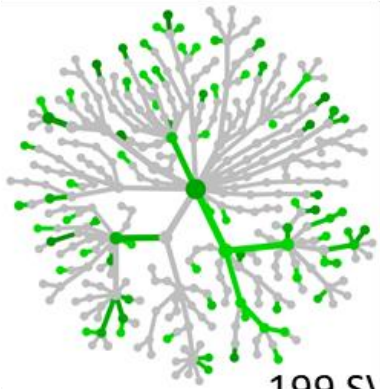
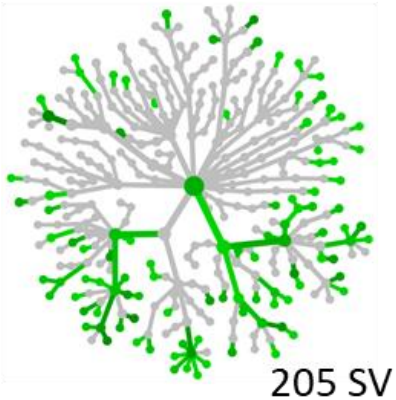
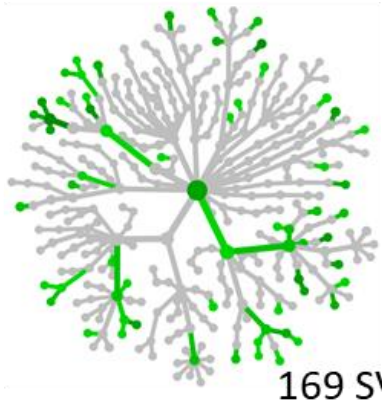
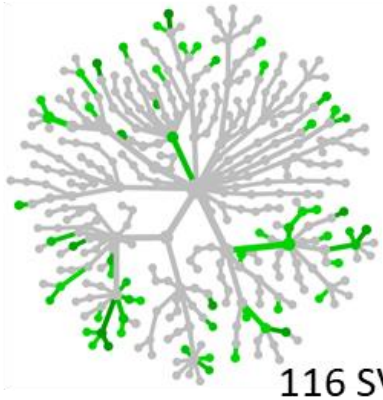


Figure 19 – Green stacked bars represent the percentage of the threshold-active sequence variants (sub-chapter 3.1.7.3 for detail) in the rhizosphere of *H. perforatum* (included in the second scoping study that were collected from the BGBM Berlin), where the active sequence variants were sorted based on the number of biological replicates where they were found to be active in. Random sampling of sequence variants for each biological replicate, with sampling size equal to the number of active sequence variants detected in that particular replicate, was conducted to see how many taxa were picked in more than 1 replicate each time random sampling was performed (out of 1000x). The result was depicted as red boxplots and only represent the non-zero values, as it is less likely that a specific taxon will be constantly picked as the number of replicates increased. N value above each boxplot depicts the number of random sampling that resulted in non-zero percentage value. None of the taxa were picked in 5 or all 6 biological replicates each time random sampling was run.

A)

	Field no. 1	Field no. 2	Greenhouse
	H. perforatum	*H. androsaemum*	*H. balearicum*
Bulk	 232 SV	 179 SV	 199 SV
Rhizosphere	 205 SV	 169 SV	 116 SV

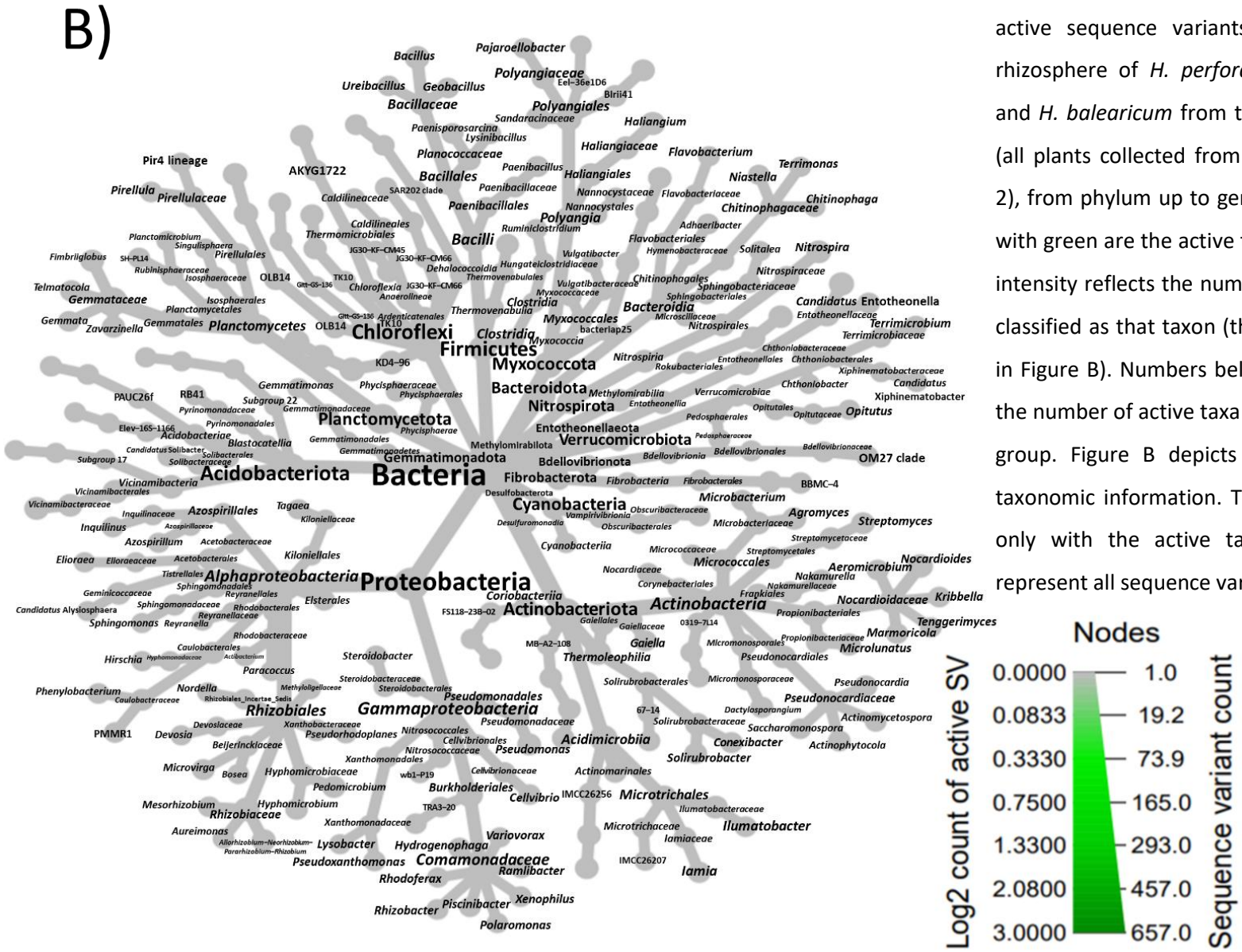


Figure 20 – Taxonomic affiliation of the group-active sequence variants in bulk soil and the rhizosphere of *H. perforatum, H. androsaemum,* and *H. balearicum* from the second scoping study (all plants collected from the BGBM Berlin; Table 2), from phylum up to genus level (A). Highlighted with green are the active taxa of each group. Color intensity reflects the number of sequence variants classified as that taxon (the numbers are depicted in Figure B). Numbers below each plot (A) depicts the number of active taxa identified in that specific group. Figure B depicts the key tree with the taxonomic information. The tree was constructed only with the active taxa and thus does not represent all sequence variants in the dataset.

4.1.2 – Investigation on *Hypericum* bacterial communities in a full-factorial crossed design-greenhouse experiment

Due to the problematic design of the first and second scoping studies that is related with insufficient number of biological replicates (only for the first scoping study) and differences that were observed even in the bulk soil bacterial communities between each plant species (for both scoping studies; see sub-chapter 4.1.1 for details), comparison of bacterial communities between different plant species could not be executed (Table 3). Thus, it is quite challenging to determine the impact of the plant species on the bacterial community of *Hypericum*. In order to tackle the issue, and to gain insights regarding the potential link between bacterial community of *Hypericum* and the content of hypericin and hyperforin, a controlled greenhouse experiment was conducted in 2019 in the BGBM Berlin (Table 3).

Table 3 – Differences between the controlled-greenhouse experiment employed in 2019 and the previous scoping studies.

Study	Plant species	Sampling year	Sampling locations	Origin	Rhizosphere samples	Sequencing platform	Biological replicates					
							Bulk soil		Rhizosphere		Roots	
							DNA	RNA	DNA	RNA	DNA	RNA
First scoping study	*H. perforatum*	2016	IPB Halle	Field	Rhizosphere only	HiSeq	2	2	3	3	0	0
	H. perforatum	2016	IPB Halle	Greenhouse	Rhizosphere only	HiSeq	0	0	3	3	0	0
	H. polyphyllum	2016	IPB Halle	Field	Rhizosphere only	HiSeq	1	1	3	3	0	0
Second scoping study	*H. perforatum*	2018	BGBM Berlin	Field	Mixed with roots	Nextseq	6	6	6	6	Mixed with rhizosphere	
	H. androsaemum	2018	BGBM Berlin	Field	Mixed with roots	Nextseq	6	6	6	6		
	H. balearicum	2018	BGBM Berlin	Greenhouse	Mixed with roots	Nextseq	5	5	5	5		
A controlled greenhouse experiment	*H. perforatum*	2019	BGBM Berlin	Greenhouse	Analysed separately	Nextseq	5	5	5	5	5	5
	H. olympicum	2019	BGBM Berlin	Greenhouse	Analysed separately	Nextseq	5	5	5	5	5	5
	H. balearicum	2019	BGBM Berlin	Greenhouse	Analysed separately	Nextseq	5	5	5	5	5	5

In this controlled-greenhouse experiment, the total (DNA-based) and active (RNA-based) bacterial communities from inside roots and the rhizosphere of *H. perforatum*, *H. olympicum*, and *H. balearicum* were investigated along with the corresponding bulk soil. A total of 5 biological replicates were investigated for each plant species. The impact of soil substrate, plant species, and habitat type (bulk soil, rhizosphere, or roots) on the total (DNA-based) and active (RNA-based) bacterial communities were determined. Active taxa for each group of bulk soil, the rhizosphere and inside the root bacterial communities were identified using a specific threshold based on the rRNA:rDNA ratio (see sub-chapter 3.1.7.3 and Figure 17).

Although plants were cultivated in 3 different soil substrates (neutral, acidic, and alkaline), only growth in acidic and alkaline soils were investigated since they delivered the most contrasting yield on plant biomass (Figure 21A). All three plant species grew poorly on alkaline substrate when compared to acidic substrate. The impact of soil substrate (acidic vs alkaline substrate) on the biomass yield seemed to be stronger for *H. perforatum* (Figure 21A). However, the fold change of

acidic and alkaline biomass was not significantly between different plant species (Figure 21B), implying similar effect of soil substrate on the plant biomass for each plant species. Moreover, the root density of *H. perforatum* on acidic substrate was the highest while the lowest was observed in *H. balearicum* (Figure 22). The same pattern was observed on alkaline substrate, albeit not shown in the picture.

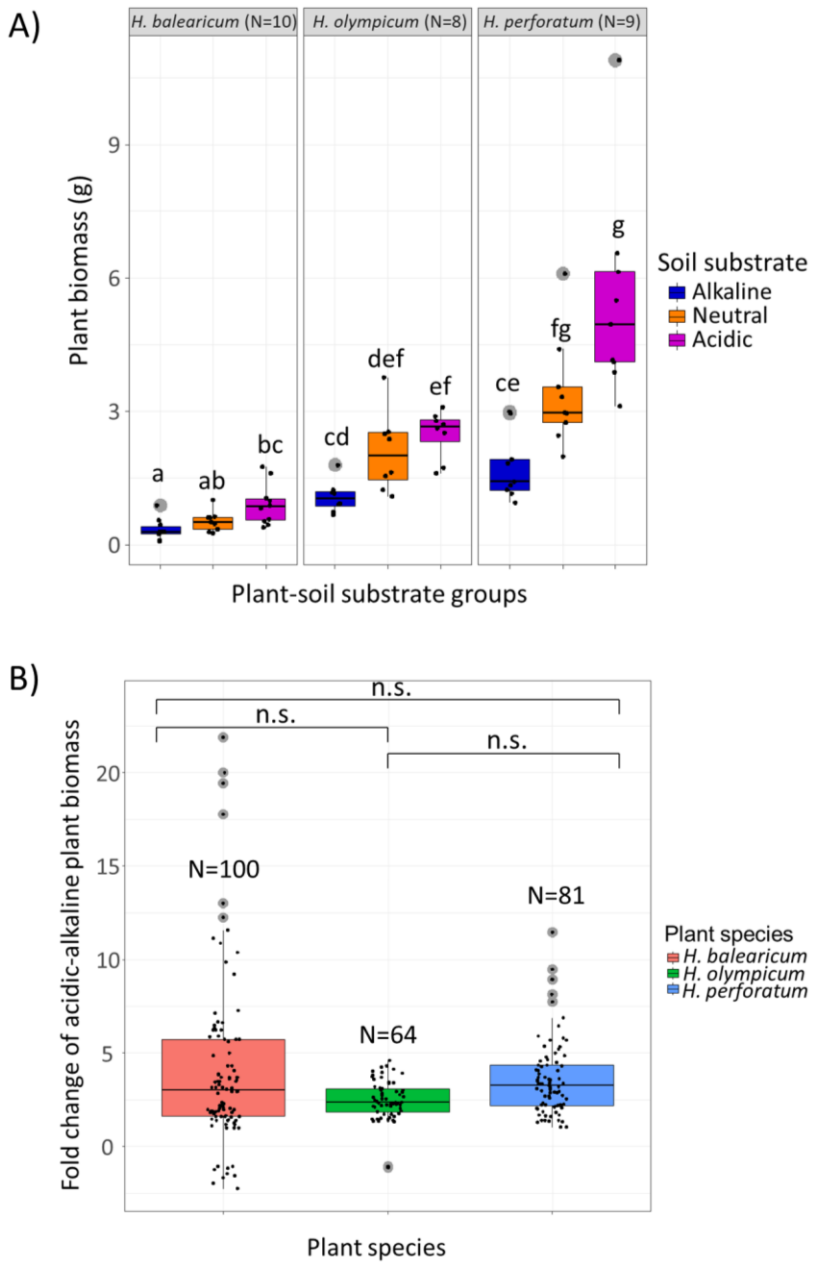


Figure 21 – The effect of soil substrate on plant biomass (A), in which letter denotes significant differences (multcomp, $p < 0.05$) between different soil substrate and plant species and grey outline depicts an outlier. B) Significant differences in fold change between acidic and alkaline biomass across distinct plant species (t-test, $p < 0.05$). N represents the number of samples for each group.

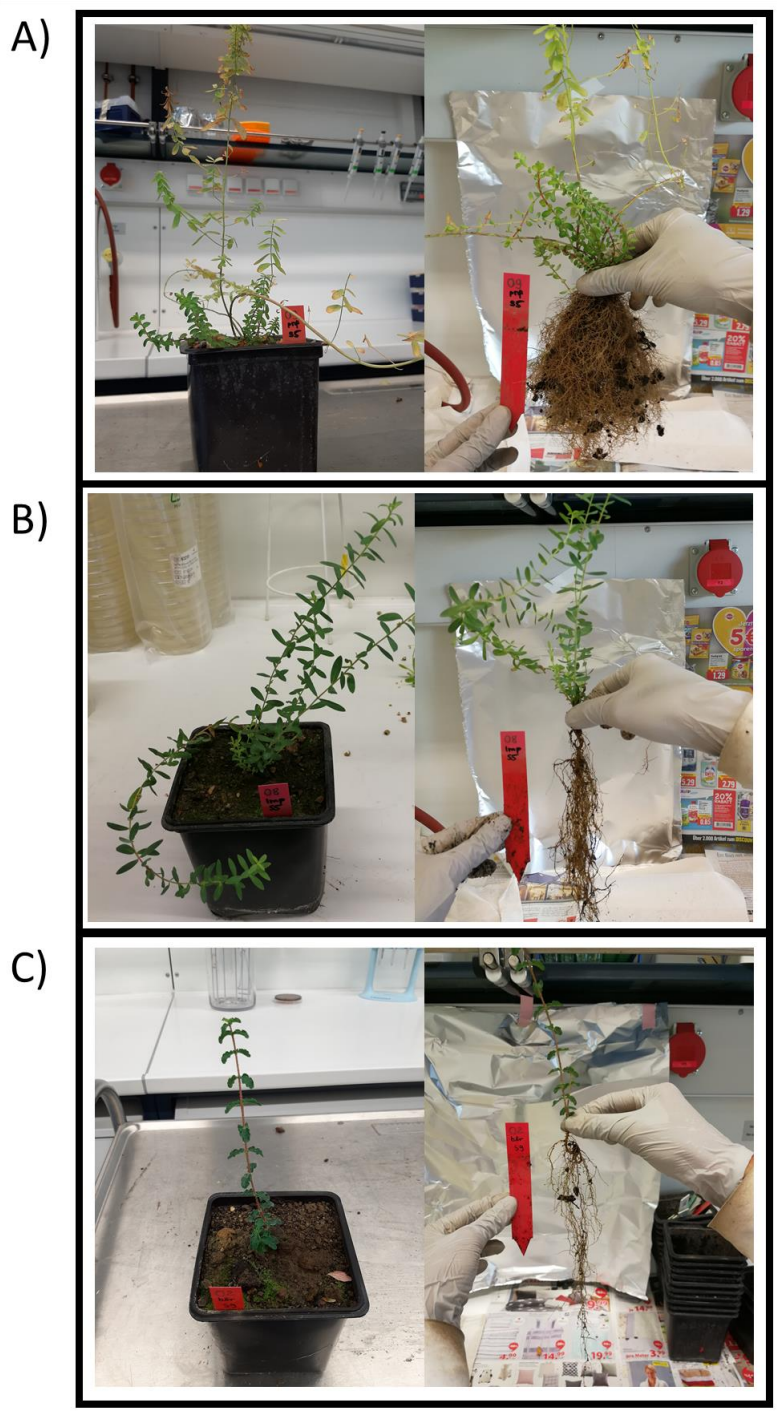


Figure 22 – *H. perforatum* (A), *H. olympicum* (B), and *H. balearicum* (C) on acidic soil substrate, several days before (left) and during the harvest (right).

In total, 83,822,290 reads which belonged to 97,175 sequence variants were retrieved after processing with Qiime2. Rarefaction curves suggested that our sequences cover most of the taxa present in the samples (Figure 23) with samples coverage estimates values ranged from 89.5% to 99.6%. To allow fair comparison between samples, the data was rarified to 58,593 reads per sample.

This left the dataset with 91,411 unique sequence variants that were distributed across 173 samples with samples coverage estimates of above 95% for all samples.

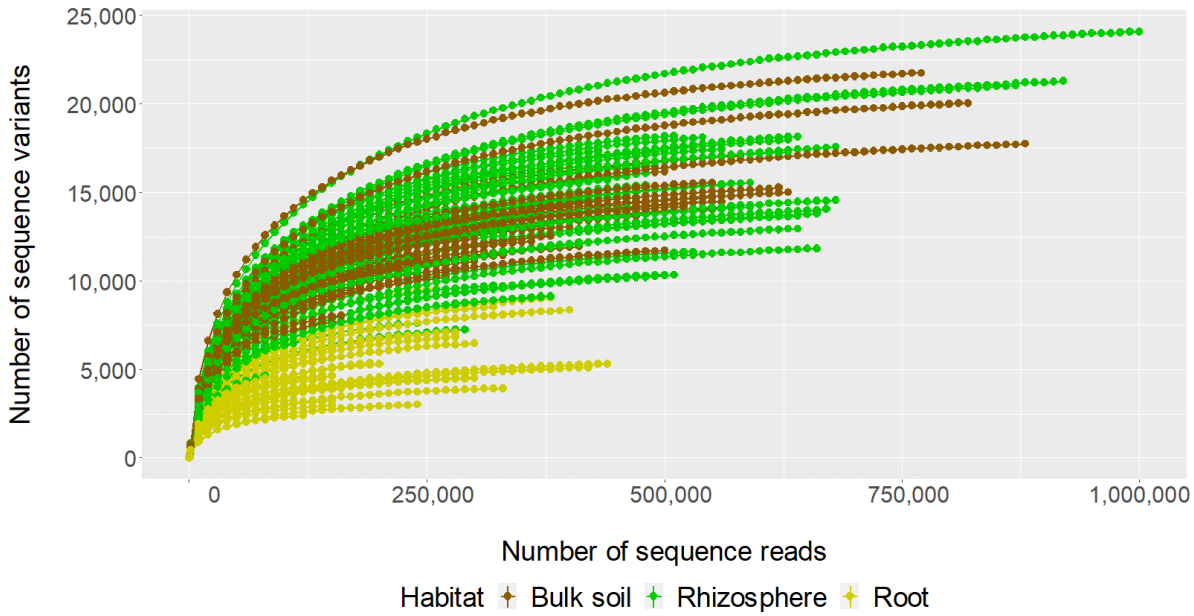


Figure 23 – Rarefaction curves for bulk soil, rhizosphere, and root samples from the controlled greenhouse experiment at sequence variant level.

To determine if distinct habitat type, plant species, and soil substrate have an effect on the diversity of bacterial communities of *Hypericum*, alpha diversity measures including richness, Shannon diversity, evenness, and alpha gambin (see 3.1.7.3 for detail description of alpha gambin) were calculated across all samples in the controlled greenhouse experiment at sequence variant level. The result revealed that the roots had lower bacterial diversity measures (richness, Shannon, and evenness; Figure 24) compared to the rhizosphere and bulk soil, regardless of the soil substrate and plant species. Lower bacterial diversity inside the roots compared to the rhizosphere is commonly known, which is related to the requirements to colonize the root and survival inside the living tissue.

Taxa richness and Shannon diversity of the total (DNA-based) and active (RNA-based) bacterial communities were similar for bulk soil samples, regardless of soil substrate and plant species, suggesting that neither soil substrate or plant species have an effect on taxa richness and diversity of bulk soil bacterial communities. Of rhizosphere samples, those belonged to *H. balearicum*-alkaline group had lower species richness, Shannon diversity and evenness compared to the other groups. Since the decrease in bacterial diversity was only associated to *H. balearicum,* this may imply the incapability of the plant species to maintain the rhizosphere bacterial diversity, as also supported by the fact that the plant grew poorly on alkaline substrate.

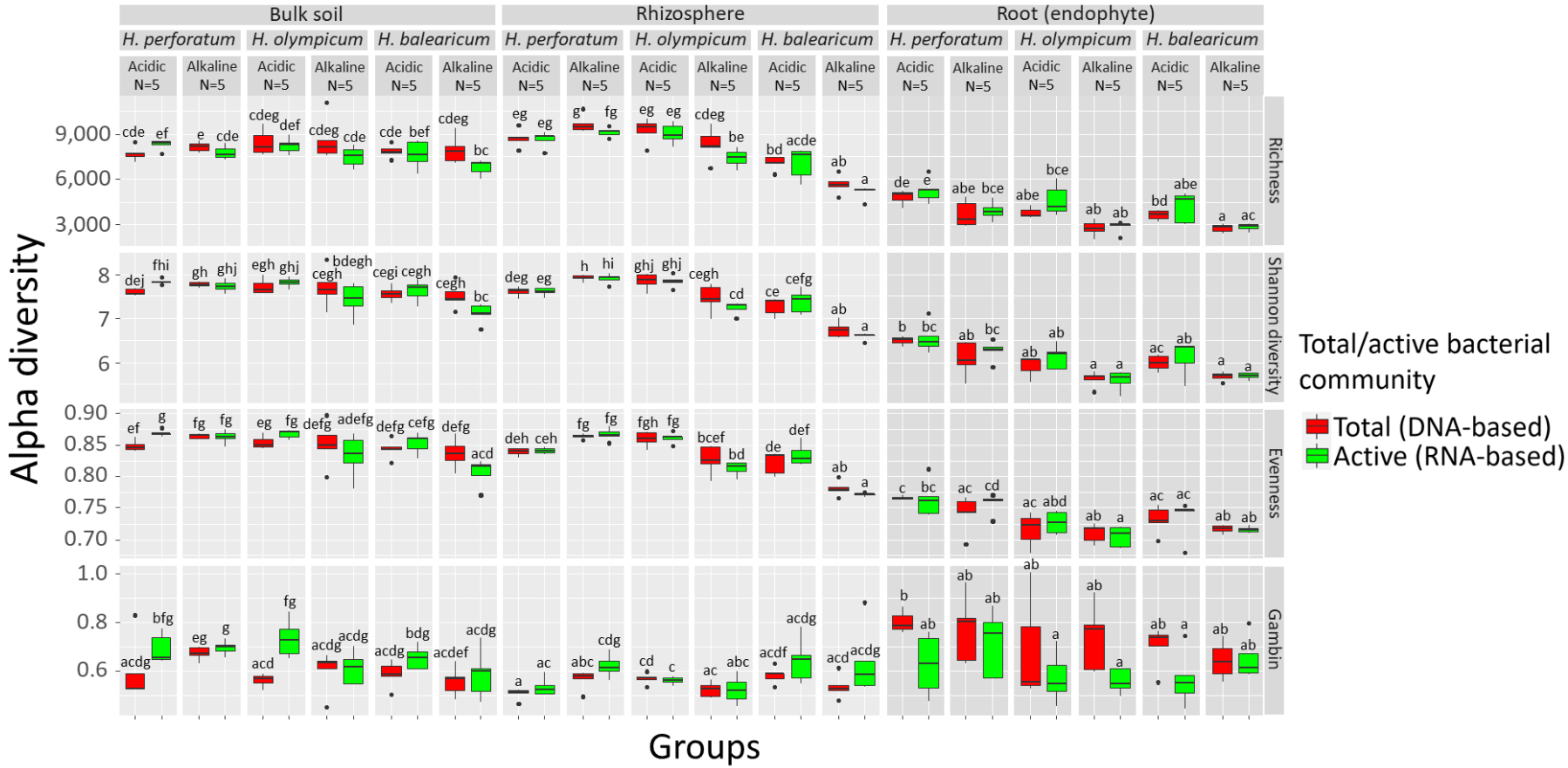


Figure 24 – Alpha diversity metrics (richness, Shannon diversity, evenness, and alpha gambin) for bacterial communities in the controlled greenhouse experiment at sequence variant level. Both total (DNA-based; red boxplots) and active (RNA-based; green boxplots) bacterial communities were investigated. For bulk soil and rhizosphere habitats, significant differences in each diversity metric between distinct soil substrate (acidic and alkaline), plant species (*H. perforatum, H. olympicum,* and *H. balearicum*), total (DNA-based) and active (RNA-based) bacterial communities, and between the two habitats are denoted with letters at the top of each box plot ($p < 0.05$, multcomp test). The test was performed separately for root habitat with the same variables tested ($p < 0.05$, multcomp test). N represents the number of samples for each group.

Almost no significant differences could be observed between bulk soil and the rhizosphere, except for the decrease of species richness and Shannon diversity in the alkaline-rhizosphere group of *H. balearicum* and the increase of species richness in the alkaline-rhizosphere group of *H. perforatum* when compared to the respective bulk soil. Since both plant species were observed to grew poorly on alkaline substrate, the contradictory results on the rhizosphere bacterial diversity may reflect the different responses of different plants to soil alkaline condition. For example, in the case of *H. perforatum,* the plant may release some compounds (via root exudation) that promote bacterial diversity in the rhizosphere in order to assist the plant in mitigating the effect of alkaline condition.

Moreover, although evenness of the root bacterial communities implies a stronger dominance based on its comparatively low values when compared to bulk soil and the rhizosphere, the dominance was not confirmed as the alpha gambin values were similar between the three habitat types (see subchapter 4.1.1; Figure 11). The result implies that instead of a small number of taxa having higher contribution to the ecological functions associated to their own habitat, diverse bacterial taxa in bulk soil, the rhizosphere, and roots are likely involved in the soil or plant-related ecological processes.

Variance partitioning analysis was conducted to measure the contribution of habitat type (bulk soil, rhizosphere and roots), soil substrate (acidic and alkaline), plant species (*H. perforatum, H. olympicum,* and *H. balearicum*), and community type (total/DNA-based and active/RNA-based communities) to the differences on the bacterial communities (Figure 25). Habitat type was the strongest driver of bacterial community, explaining approximately 23.6% of variance. The second highest variance (10.3%) could be explained by differences between total (DNA-based) and active (RNA-based) bacterial communities. This was followed by soil substrate and plant species, each contributing to 7.7% and 6.4% of variance, respectively. All values were statistically significant (ANOVA, $p < 0.01$).

NMDS plot based on weighted UniFrac distances was constructed to further evaluate the drivers of bacterial community structure and composition. Stress value of NMDS indicated a good-fit ordination in 3 dimensions that is not based upon random process. Grouping according to habitat type is evident along the first axis of NMDS (Figure 26A; PERMANOVA, $r2 = 0.24$, $p < 0.001$), supporting the result of variance partitioning analysis. The scale of differences between bulk soil and roots (endophyte) communities was the largest according to the grouping on the ordination plot (located far from each other while rhizosphere samples were in between). Separation between the total (DNA-based) and active (RNA-based) bacterial communities could be observed on the NMDS plot (Figure 26A), but it was more pronounced when the second axis was plotted against the third one (Figure 26B; PERMANOVA, $r2 = 0.10$, $p < 0.001$). Separation of samples based on soil substrate

and plant species was less pronounced on the ordination plot but this may reflect the small variances (7.7% and 6.4% of variance, respectively) that could be explained by those two varibles on bacterial community structure and composition based on variance partitioning analysis.

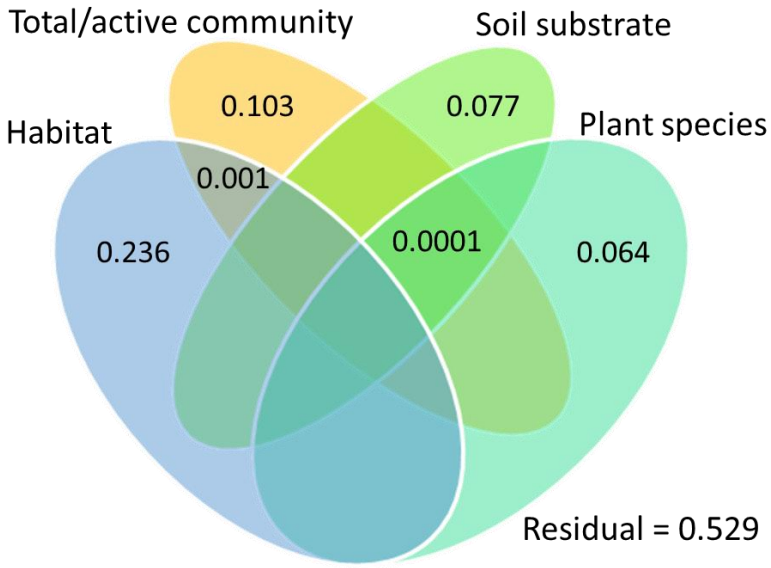


Figure 25 – Variance partitioning analysis based on weighted UniFrac distances at sequence variant level, depicting the proportion of variance of bacterial community composition that can be explained by habitat type (bulk soil, rhizosphere, and roots), differences between total (DNA-based) and active (RNA-based) bacterial communities, soil substrate (acidic and alkaline), and plant species (*H. perforatum, H. olympicum,* and *H. balearicum*) in the controlled greenhouse experiment.

Based on variance partitioning analysis, we concluded that the habitat type is the strongest driver of bacterial communities associated to *Hypericum*. In order to determine the contribution of other drivers (soil substrate, plant species, and community type) to the structure and composition of bacterial communities in each habitat, we performed variance partitioning for each bulk soil, rhizosphere, and root bacterial communities (Figure 27). All individual and shared fractions were tested with ANOVA and all were found to be significant ($p < 0.05$). Soil substrate governed bulk soil bacterial communities, explaining approximately 30% of variance. Soil substrate effect was also evident on the rhizosphere and root bacterial communities, albeit less pronounced (12% and 8%, respectively). On the other hand, plant species impact was similar on both rhizosphere and root bacterial communities, with approximately 20% of variance explained, while the impact on bulk soil communities was small (2% of variance explained). Differences between the total (DNA-based) and active (RNA-based) bacterial communities were also observed regardless of habitat type, which imply differences in the transcriptional activity among resident bacterial taxa.

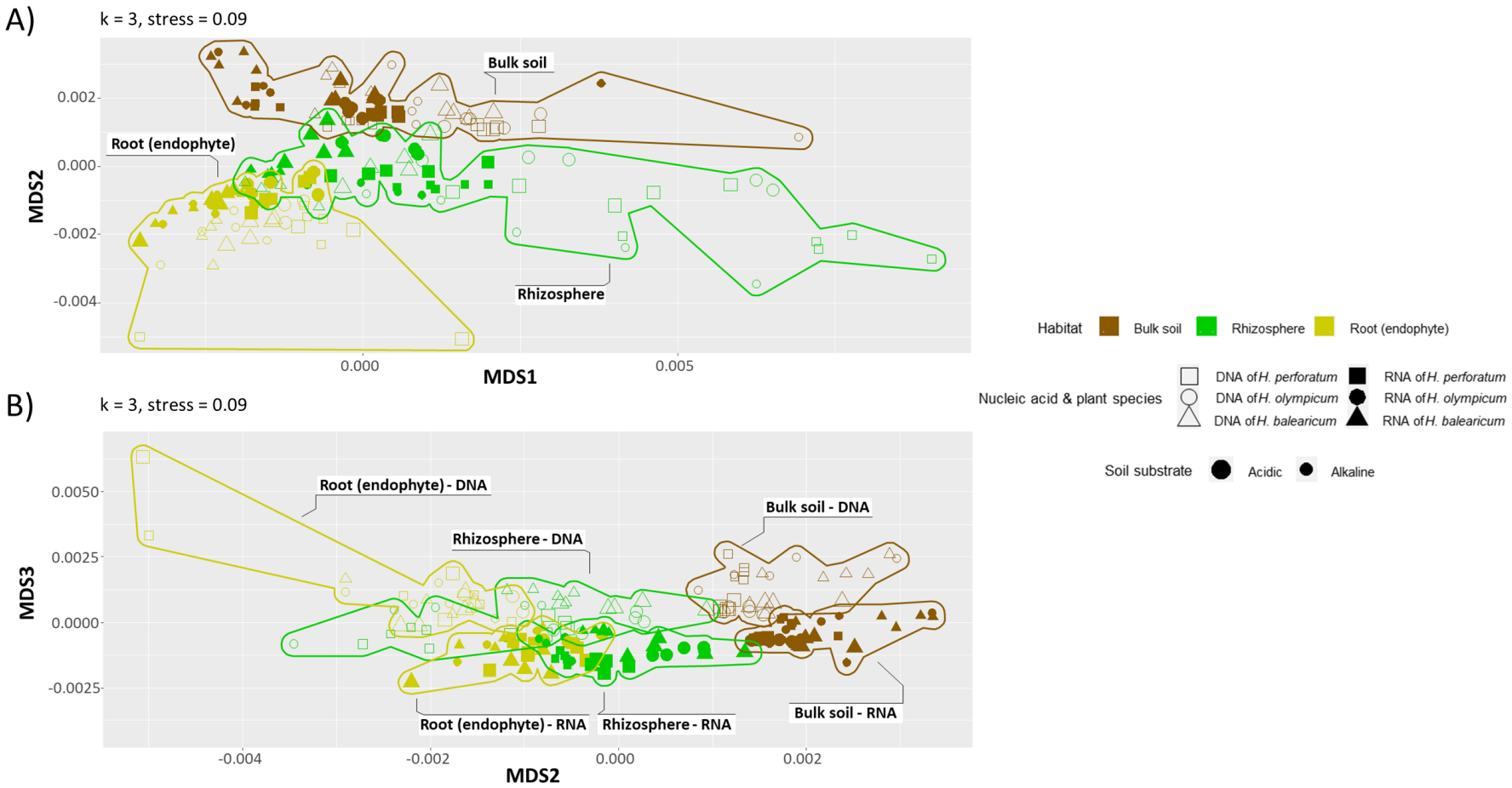


Figure 26 – Non-metric multidimensional scaling based on weighted UniFrac distances at sequence variant level, along the first and second axis (A; emphasizing differences of bacterial communities structure and composition due to different habitat types; PERMANOVA, $r^2 = 0.24$, $p < 0.01$) and along the second and third axis (B; emphasizing differences between the total (DNA-based) and active (RNA-based) bacterial communities; PERMANOVA, r2 = 0.10, $p < 0.001$).

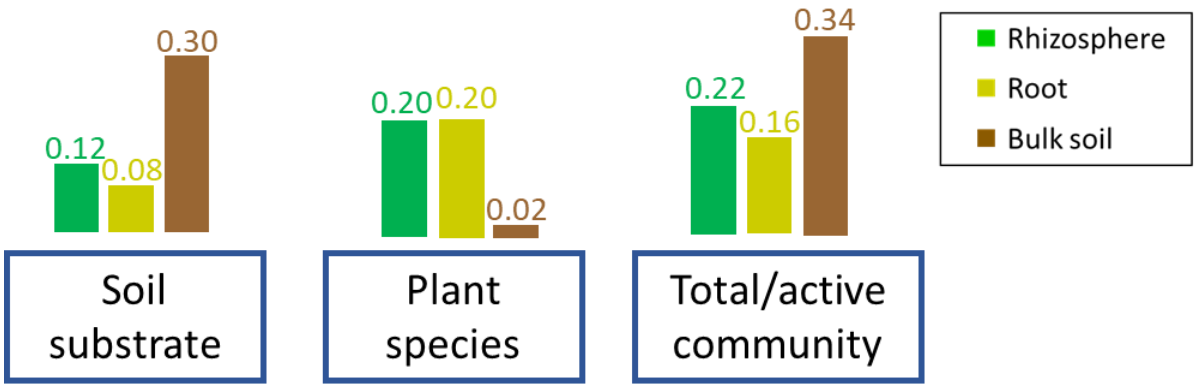


Figure 27 – Variance partitioning analysis depicting the proportion of variance of bulk soil, rhizosphere, and root bacterial communities (at sequence variant level) that can be explained by soil substrate, plant species, and differences between the total (DNA-based) and active (RNA-based) bacterial communities. For all three habitats, approximately 1% of variance can be explained by shared effect of all three variables. Residual values of variance partitioning for the rhizosphere, roots, and bulk soil were 49.1%, 57.3%, and 36.8%, respectively.

In order to evalute further the drivers of bulk soil, rhizosphere, and root (endophytic) bacterial communities, NMDS plot was constructed based on weighted UniFrac distances for each habitat. The ordination of the three NDMS plots were considered to be fair (less likely based on random process with stress values ≤ 0.1). NMDS plot of bulk soil bacterial communities revealed a clear separation based on soil substrate (Figure 28) and thus confirmed the impact of soil substrate on bulk soil bacterial communities based on variance partitioning analysis (30 % of variance explained; Figure 27). In addition, the total (DNA-based) and active (RNA-based) communities were also clustered separately on the plot, implying different level of metabolic activities across soil bacteria. The separation based on plant species, which explains smaller amount of variance (around 2%), could not be observed even when the third axis was included on the ordination plot (not shown). These results confirmed major impact of soil substrate on bulk soil bacterial communities whereas plant species impact was less pronounced.

NMDS plot of the rhizosphere bacterial communities revealed a clear separation based on soil substrate and plant species (Figure 29A). In addition, the total (DNA-based) and active (RNA-based) bacterial communities were also clustered separately on the plot, implying different level of metabolic activities across rhizospheric bacteria. The impact of all variables to the rhizosphere bacterial communities were significant (PERMANOVA, p value < 0.001). Although samples from total (DNA-based) and active (RNA-based) communities seemed to overlap when using only two dimensions on NMDS, a clear separation was observed when the third axis was added (Figure 29B),

confirming the result of variance partitioning. The impact of plant species on the rhizosphere bacterial communities was more pronounced on alkaline substrates on the ordination plot (Figure 29A), which may reflect a different survival mechanism of each plant species to endure alkaline condition since all plants grew poorly on alkaline soil. The impact of soil substrate on the rhizosphere bacterial communities is evident on the ordination plot, as acidic and alkaline samples are separated along the second MDS axis (Figure 29A). These results confirmed the impact of soil substrate and plant species on the rhizosphere bacterial communities.

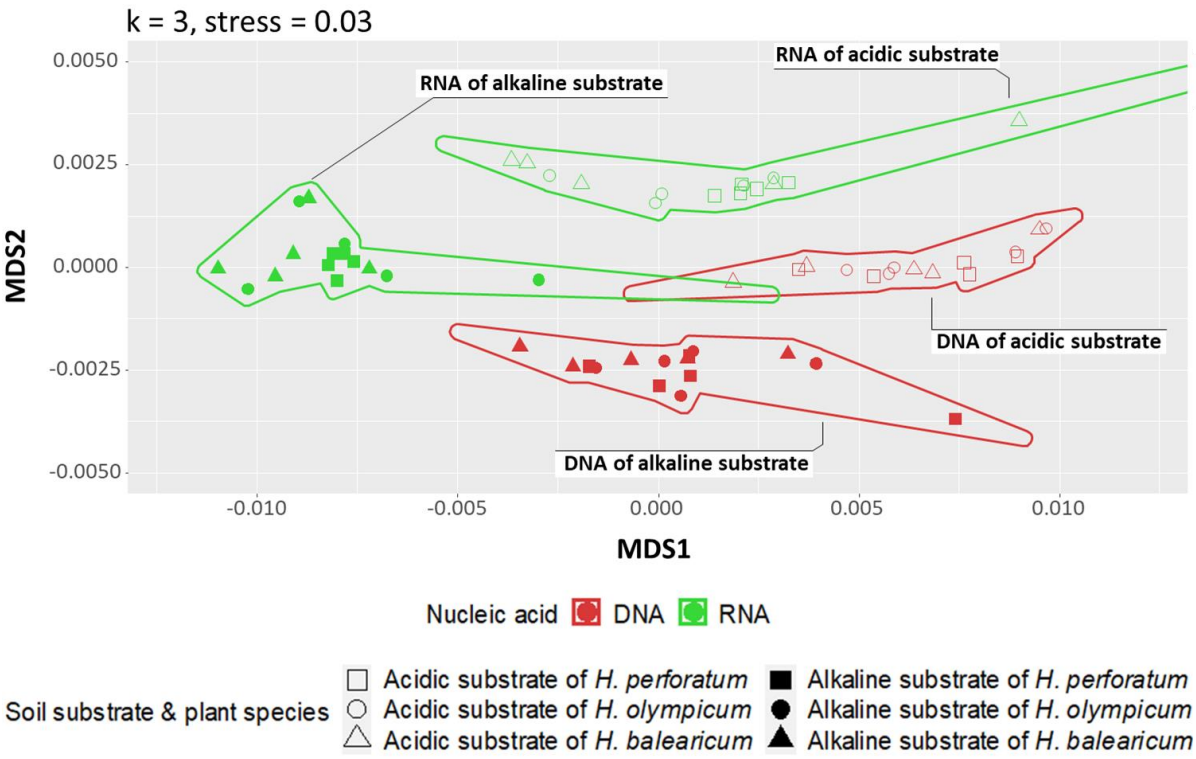


Figure 28 – NMDS plot based on weighted UniFrac distances at sequence variant level of the total (DNA-based) and active (RNA-based) bacterial communities in bulk soil.

NMDS plot of the root (endophytic) bacterial communities revealed separation based on plant species (Figure 30), supporting the result of variance partitioning (20% of variance explained; Figure 27). As had been observed in bulk soil and the rhizosphere, separation of the total (DNA-based) and active (RNA-based) bacterial communities was also evident inside the roots, confirming differences in the transcriptional activity among bacteria inside the roots (16% of variance explained; Figure 27). The separation based on soil substrate was also observed but to a smaller degree (Figure 30), supporting the result of variance partitioning where soil substrate only explained 8% of variance. The effect of all three variables was significant (PERMANOVA, $p < 0.001$). These results confirmed the impact of plant species and soil substrate on the root (endophytic) bacterial communities.

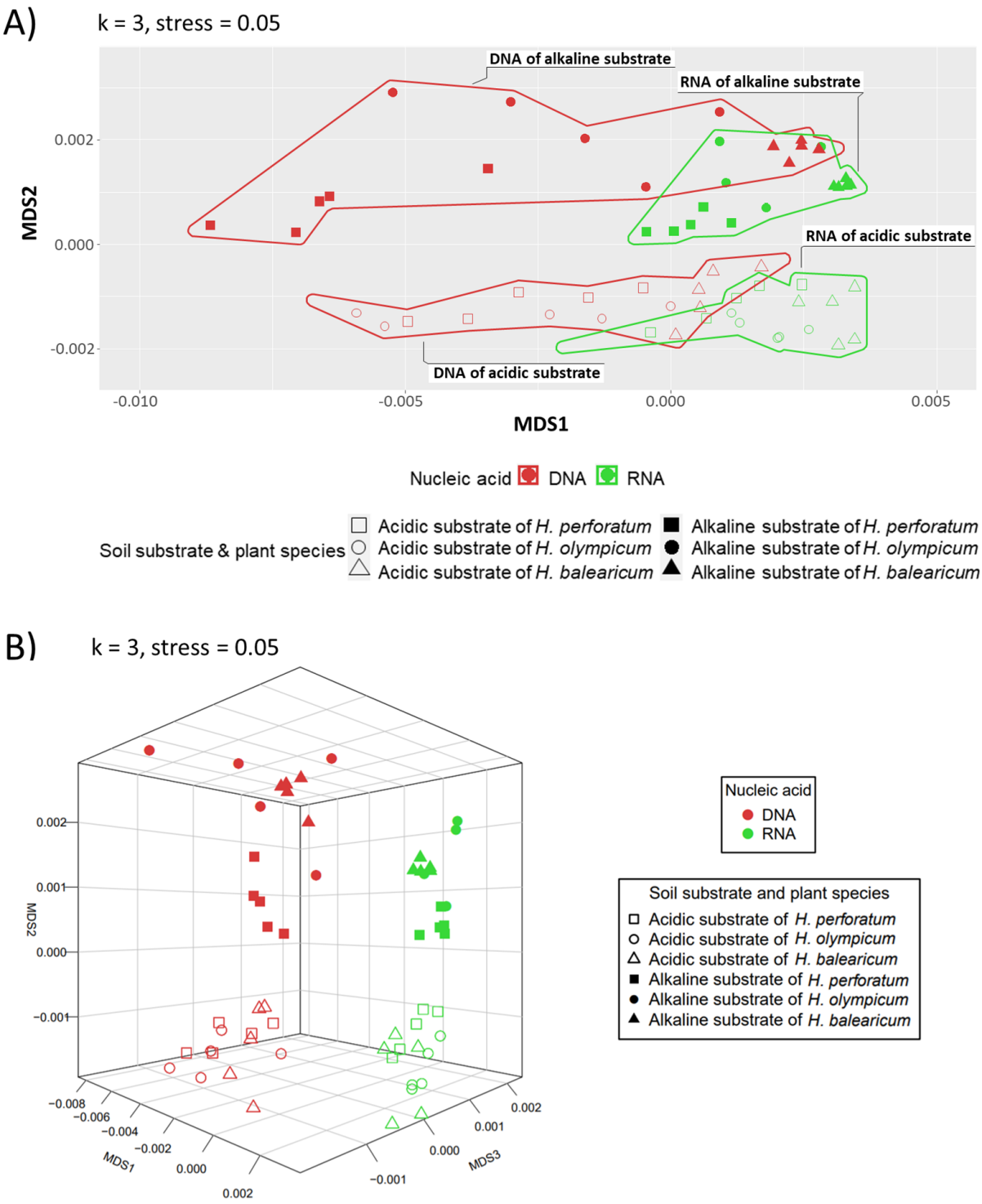


Figure 29 – NMDS plot based on weighted UniFrac distances at sequence variant level of the total (DNA-based) and active (RNA-based) bacterial communities in the rhizosphere along the first and second axes (A), or with addition of the third axis (B).

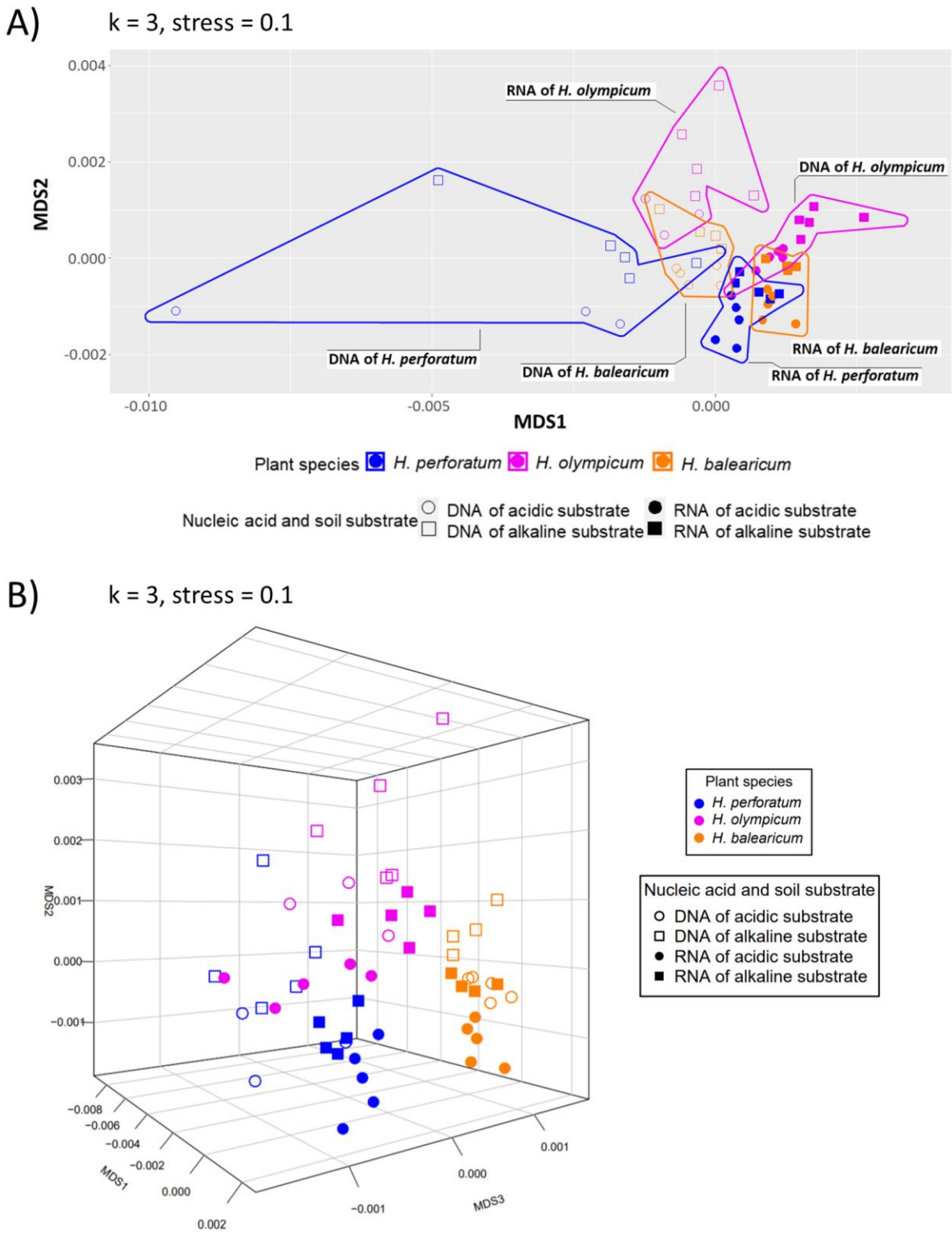


Figure 30 – NMDS plot based on weighted UniFrac distances at sequence variant level of the total (DNA-based) and active (RNA-based) bacterial communities inside the roots along the first axis (A), or with addition of the third axis (B).

To determine bacterial taxa (up to genus level) that may actively contribute to the ecological functions in their respective habitat, comparison between the total (DNA-based) and active (RNA-based) bacterial communities was conducted for each habitat using the R package metacoder (Figure 31). Enriched taxa in the active communities (RNA-based) may perform important functions related to soil ecosystem functioning in bulk soil or related to the fitness of the host plant in the rhizosphere and/or inside the roots. The result revealed that the enriched taxa between total (DNA-based) and active (RNA-based) bacterial communities were similar for bulk soil and the rhizosphere, suggesting that the bacterial communities may be more similar between those two habitats compared to the roots. Less taxa were enriched inside the roots, but this is probably due to the lower richness verified inside the roots (Figure 24).

At phylum level, the abundance of Planctomycetota, Myxococcota, WPS-2, Entotheonellaeota, Hydrogenedentes, and Deferrisomatota was significantly higher in the active (RNA-based) bacterial communities when compared to the total ones (DNA-based), independently of habitat type (Figure 31). At genus level, the enriched taxa in the active (RNA-based) communities included *Piscinibacter, Phenylobacterium*, *Polyangium, Pedomicrobium, Haliangium, Phaselicystis,* OM27 clade of *Bdellovibrionaceae*, *Pirellula, Gemmata, Fimbriiglobus,* OLB14 of *Anaerolineales, Telmatocola, Sporichthya, Marinoscillum, Zavarzinella, Solirubrobacter, Planctopirus, Nannocystis, Anaeromyxobacter, Pajaroellobacter,* "*Candidatus* Entotheonella", *Bryobacter*, "*Candidatus* Solibacter", *Sandaracinus, Ramlibacter, Pseudenhygromyxa, Deferrisoma,* Pir3 lineage of *Planctomycetaceae*, *Plaudibaculum, Desulfonatronum,* and *Tuwongella.* This result may reflect the ability of these taxa to utilize a wide range of substrate and thus enriched in the active communities (RNA-based) of all habitats. These taxa are probably not performing specific functions affiliated with a particular habitat.

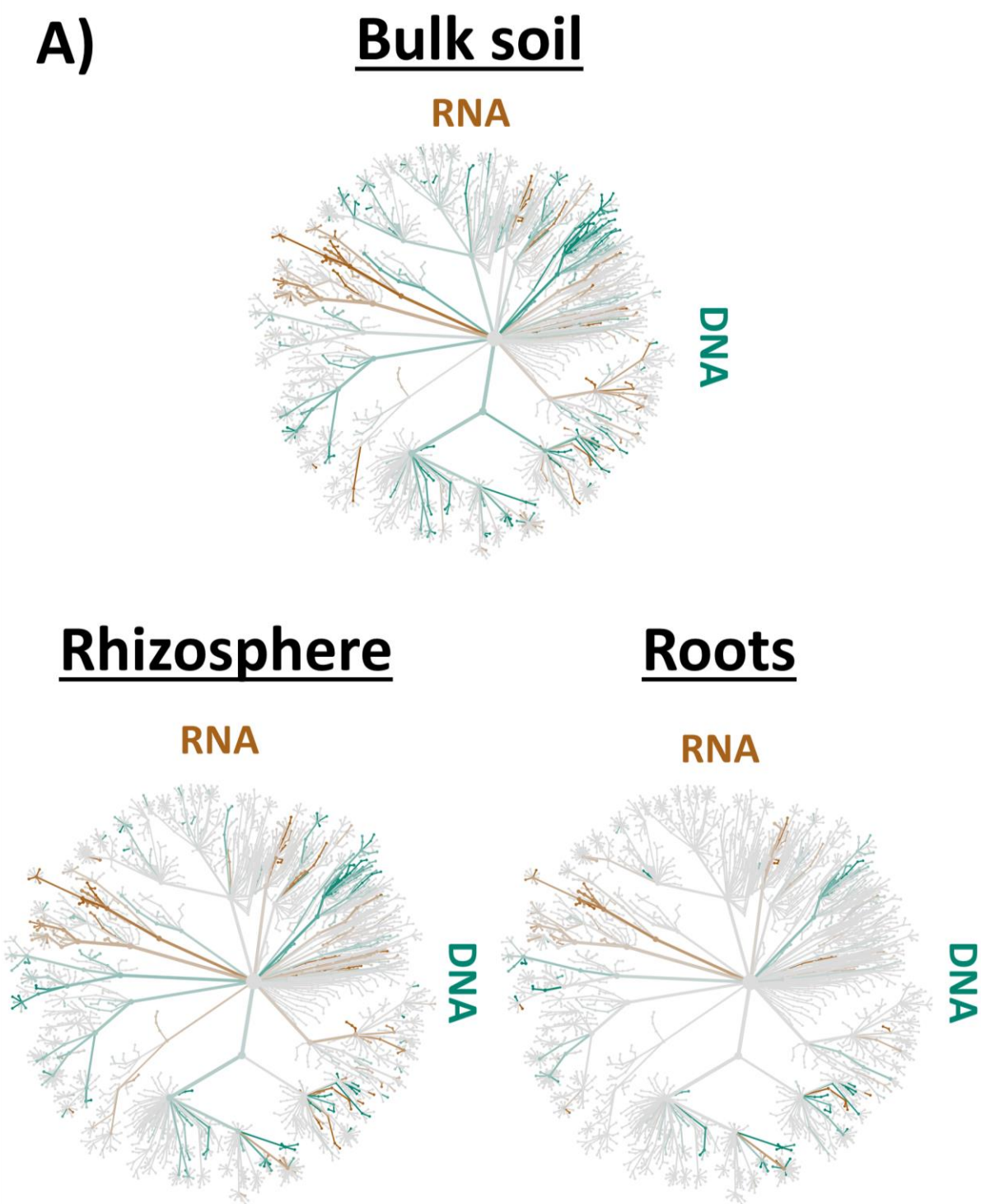
On the other hand, Verrucomicrobiota Patescibacteria, FCPU426, Elusimicrobiota, and Dependentiae along with some genera such as *Methylotenera, Hydrogenophaga, Acidibacter*, *Pseudolabrys, Dongia, Luteolibacter,* UTCFX1 of *Anaerolinaceae*, DSSD61 of *Nitrosomonadaceae, Terrimonas, Polycyclovorans, Estrella, Laceyella,* OLB14 of *Anaerolineae, Planifilum, Albirhodobacter, Bdellovibrio, Hydrogenispora,* UBA6140 of *Methylophilaceae, Aquicella, Mycoplasma, Methylophaga, Neochlamydia,* JGI 0001001-H03 of *Blastocatellaceae,* "*Candidatus* Rubidus", *Vampirovibrio, Sulfurifustis, Legionella, Caldicoprobacter,* "*Candidatus* Omnitrophus", *Tumebacillus, Haloferula,* TM7x of *Saccharimonadaceae*, and *Haloplasma* were consistently enriched in total (DNA-based) bacterial communities when compared to active ones (RNA-based) regardless of habitat type. Thus, these taxa may generally dormant or represent relic DNA across all habitats.

Of bulk soil bacterial communities, the taxa that were enriched in the active (RNA-based) communities included *Rhizobacter, Pelomonas, Devosia, Rubrivivax, Vulgatibacter, Tagaea, Brevundimonas, Halanaerobium, Corallococcus, Sphingobacterium, Thermosediminibacter* and *Azoarcus.* These taxa may actively contribute to soil ecosystem functioning. Moreover, Actinobacteriota, Bdellovibrionota, Desulfobacterora, *Bosea, Steroidobacter, Pseudonocardia, Azospirillum,* "*Candidatus* Chloroploca", *Dactylosporangium, Lautropia, Rhodopirellula, Xenophilus Archangium, Luteitalea,* IMCC26207 of *Microtrichaceae*, *Blastopirellula, Singulisphaera, Vicinamibacter, Oikopleura, Nitrosomonas, Schlesneria, Conexibacter, Ottowia,* and *Rubrivirga* were enriched in the active (RNA-based) communities in both bulk soil and the rhizosphere. Although they may be active in the rhizosphere, their ecological roles are probably not specifically related to the plant species since they are also potentially active in bulk soil habitat.

Reyranella, Nodosilinea PCC-7104, *Tychonema* CCAP 1459-11B, *Euzebya, Leptolyngbya* ANT.L52.2*, Ornithinimicrobium, Nitrospira, Paludisphaera, Elioraea, Actinomycetospora,* and *Geoalkalibacter* were among the enriched taxa in the active (RNA-based) bacterial communities only in the rhizosphere whereas PMMR1 of *Caulobacteraceae*, *Opitutus,* subgroup 17 of Acidobacteriota, *Ilumatobacter*, *Rhizocola, Algisphaera, Inquilinus, Methyloversatilis, Emticicia,* and *Rubinisphaera* were the enriched taxa in the active (RNA-based) bacterial communities only inside the roots*.* Moreover, Acidobacteriota, WS2, *Bacillus,* CL500-29 marine group of *Ilumatobacteraceae, Paenibacillus, Byssovorax, Gimesia,* and *Phaeodactylibacter* were enriched in the active (RNA-based) communities in the rhizosphere and inside the roots. These taxa may perform specific functions that are important for the host plant in general. In addition, the taxa residing inside the roots may have the ability to colonize the roots and adapt to the living condition inside the living tissues.

Of the enriched taxa in the total (DNA-based) bacterial communities when compared to active ones (RNA-based), *Niastella, Curvibacter, Solimonas, Peredibacter, Aridibacter, Rhodobacter, Bythopirellula, Sphaerobacter,* "*Candidatus* Protochlamydia", *Algoriphagus,* and *Bacteriovorax* were specifically enriched inside the roots whereas *Opitutus,* IMCC26134 of *Opitutaceae, Diplosphaera, Pseudoxanthomonas, Pedosphaeraceae, Alterococcus, Dyadobacter, Erysipelatoclostridiun, Imperialibacter, Marmoricola, Ereboglobus, Kinneretia, Pedosphaera, Halocella, Crocinitomix, Subsaxibacter, Kroppenstedtia, Ignavibacterium, Pleomorphomonas,* "*Candidatus* Methylopumilus", *Defluviitalea, Kazania, Corynebacterium, Melghirimyces, Metallibacterium, Alkaliphilus, Effusibacillus,* and *Dehalobacterium* were specifically enriched in the rhizosphere. In addition, *Thalassospira, Methylophilus, Lacunisphaera, Methylobacillus,* and *Thermostaphylospora* were among the enriched taxa in the total communities in the rhizosphere and inside the roots. These

results imply that these taxa, although may be found abundant in the rhizosphere and inside the roots, are not necessarily active. Thus, the result also emphasizes the importance of employing activity-based analysis (in this case RNA-based analysis) to correctly identify the key players of the habitat of interest.



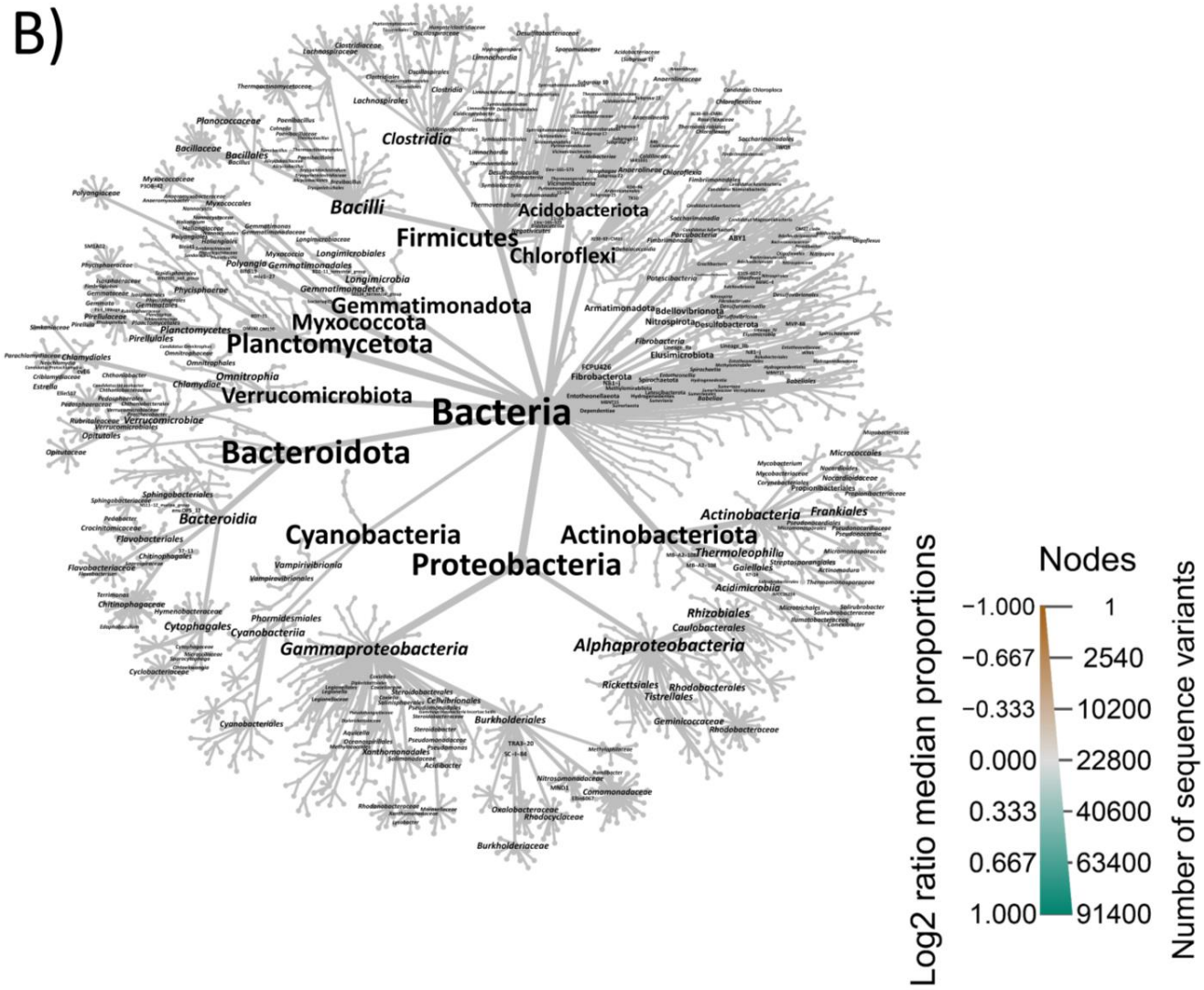


Figure 31 – Pairwise comparison of bacterial communities between the total (DNA-based) and active (RNA-based) communities of bulk soil, rhizosphere, and root samples in the controlled greenhouse experiment, from phylum up to genus level (A). Wilcoxon rank-sum test, followed by a Benjamini-Hochberg (FDR) correction for multiple comparisons, were performed to determine significant differences ($p < 0.05$). Significantly enriched taxa were colored according to the community type (either total or active bacterial communities) where they were found to be enriched in (A) and the color intensity reflects log-2 ratio in median proportions (the values are depicted in Figure B). Figure B depicts the key tree with the taxonomic information. Node diameter reflects number of sequence variant classified as that taxon whereas edge width reflects number of reads.

In order to evaluate the impact of soil substrate on bulk soil, rhizosphere, and root bacterial communities, and to determine the preferences of bacterial taxa to different soil substrates, comparison of bacterial communities between acidic and alkaline substrates was conducted for each habitat. Similar to the result of comparison between total (DNA-based) and active (RNA-based) communities, the comparison between distinct soil substrates implies higher similarity of bacterial communities between bulk soil and the rhizosphere when compared to the roots (Figure 32). Again, this may be explained by the lower diversity inside the roots compared to bulk soil and the rhizosphere.

Between distinct soil substrates, the abundance of PMMR1 of *Caulobacteraceae,* along with *Piscinibacter, Pseudomonas,* SWB02 of *Hyphomonadaceae, Niastella, Pelomonas, Thalassospira, Curvibacter, Bosea, Lacunisphaera, Inhella, Azotobacter, Stella, Methylobacillus, Noviherbaspirillum, Kineosporia, Permianibacter, Ahniella, Peredibacter, Rhodobacter,* UBA6140 of *Methylophilaceae, Vogesella,* "*Candidatus* Paracaedibacter", 11-24 of *Blastocatellaceae, Cupriavidus, Pseudoduganella,* "*Candidatus* Rubidus", *Oikopleura, Haliscomenobacter, Roseomonas, Bacteriovorax,* and *Sediminibacterium* was significantly higher in the alkaline samples, regardless of the habitat type. In addition, the alkaline-enriched taxa only in bulk soil included *Streptomyces, Methylotenera, Hydrogenophaga, Rhizobacter, Nodosilinea* PCC-7104*,* IMCC26134 of *Opitutaceae, Pseudonocardia, Limnobacter, Euzebya,* and *Leptolyngbya* ANT.L52.2. These bacterial taxa most likely prefer alkaline to acidic condition.

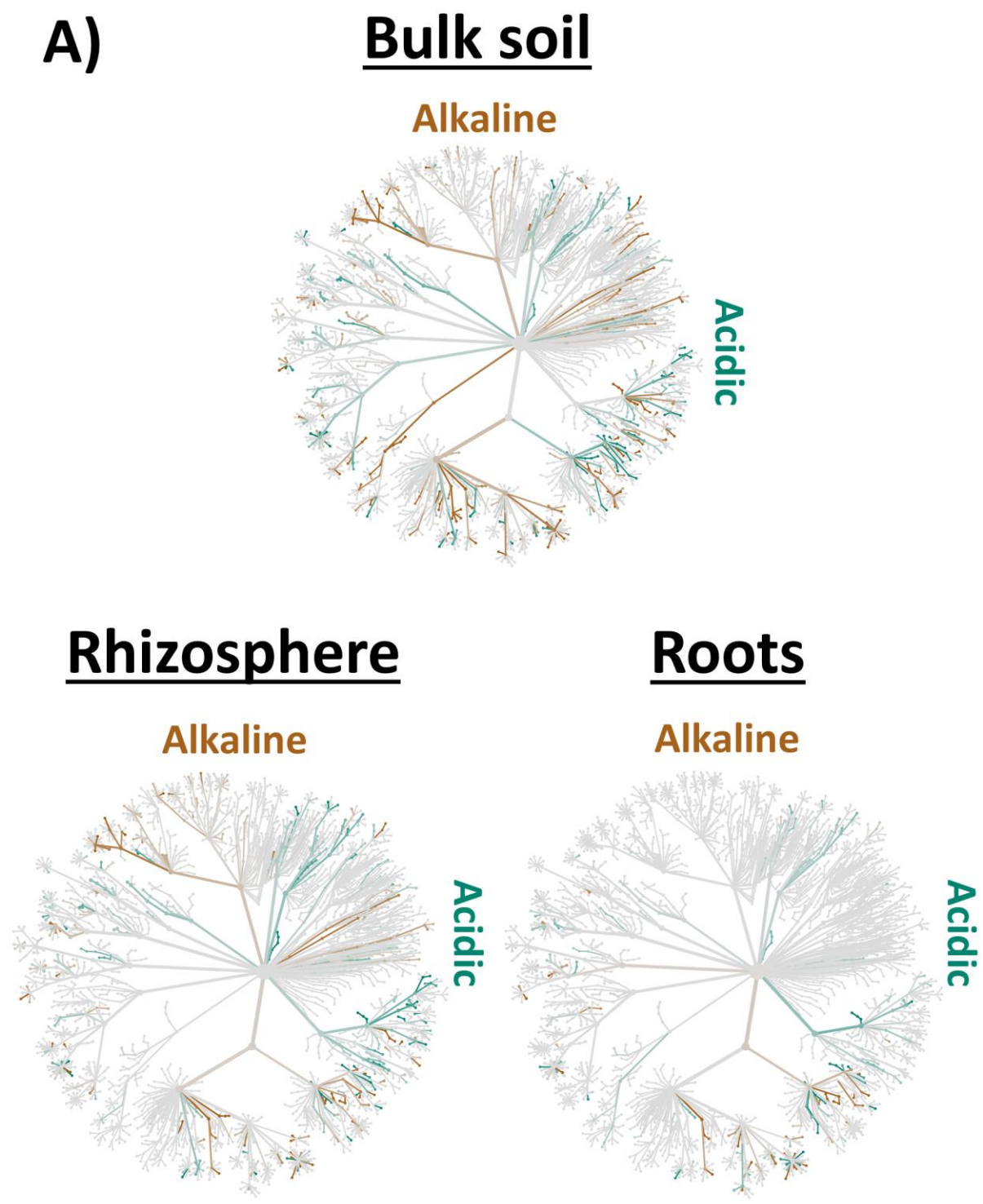
Moreover, the alkaline-enriched taxa only in the rhizosphere included *Methylophylus, Ilumatobacter, Arenibacter, Phreatobacter, Ammoniphilus, Gemmobacter, Ereboglobus, Anaerocolumna, Blastococcus, Pseudoclostridium, Anaerovorax, Ammoniibacillus, Longispora, Thermobispora, Halocella, Pseudobacteroides, Thermincola, Jahnella, Marispirillum, Kroppenstedtia, Lutispora, Sedimentibacter, Oceanobacillus, Hungateiclostridium, Novosphingobium,* "*Candidatus* Methylopumilus"*, Terrisporobacter, Anaerofustis,* R-7 group of *Christensenellaceae*, *Sinosporangium, Acetanaaerobacterium, Stenotrophobacter, Gracilibacillus, Actimicrobium, Desulfohalotomaculum, Limnohabitans,* and *Dehalobacter.* Since they were specifically enriched on alkaline substrate only in the rhizosphere and due to the fact that all plant experiencing poor growth on alkaline substrate, their enrichment may also be related to plant survival mechanism against alkaline condition.

Less taxa were found enriched on alkaline substrate specifically inside the roots. These included *Opitutus,* SM1A02 of *Phycisphaeraceae*, *Vicinamibacter,* "*Candidatus* Protochlamydia"*, Schlesneria, Sulfurifustis,* "*Candidatus* Berkiella", *Paucibacter, Fulvivirga, Methyloversatilis,* and *Cereibacter*. Since the pH of inner-root habitat is most likely constant and not affected by external soil substrate, the

enrichment of these taxa inside the roots may be caused by other factors such as nutrient availability inside the roots or their adaptation to the living condition inside the living tissues.

Firmicutes, Nitrospirota, Fibrobacterota, Spirochaetota, Zixibacteria, *Cellvibrio, Hyphomicrobium, Diplosphaera, Bacillus, Geobacillus, Lysinibacillus, Solimonas, Sporocytophaga, Coxiella, Nitrospira, Ureibacillus, Planctopirus, Tuberibacillus, Chitinimonas, Leptospira, Methylibium, Ketobacter, Planifilum, Aridibacter, Aquabacterium, Bdellovibrio, Thermomonospora, Thermopolypora, Thermoactinomyces, Bythopirellula, Clostridium* sensu stricto 7, *Turneriella, Brevibacillus, Aneurinibacillus, Cohnella, Blastospirellula, Vulgatibacter, Sphaerobacter, Thermobacillus,* JGI 0001001-H03 of *Blastocatellaceae*, and *Symbiobacterium* were among the alkaline-enriched taxa in the rhizosphere and bulk soil. Again, this may reflect the preference of these taxa to alkaline compared to acidic condition.

On the other hand, bacterial taxa that were enriched in acidic substrate regardless of the habitat type included Gemmatimonadota, Acidobacteriota, Chloroflexi, SAR324 clade, WS2, *Sphingomonas, Mesorhizobium, Pseudolabrys, Nocardioides, Gaiella,* OM27 clade of *Bdellovibrionaceae, Alicyclobacillus, Chhtoniobacter, Planctomicrobium, Iamia, Arthrobacter, Agromyces, Azospirillum, Polaromonas,* CL500-29 marine group of *Ilumatobacteraceae, Ferruginibacter,* "*Candidatus* Chloroploca", *Rhodanobacter, Kribbella, Aeromicrobium, Gemmatimonas, Dactylosporangium, Pedobacter, Herminiimonas, Flavisolibacter, Chryseolinea, Rhodopirellula, Archangium, Luteitalea,* RB41 of *Pyrinomonadaceae*, *Flavitalea,* "*Candidatus* Udaeobacter", BIyi10 of *Pseudohongiellaceae,* CL500-3 of *Phycisphaeraceae*, *Oligoflexus, Rhizocola, Luteimonas, Singulisphaera, Thalassobaculum,* "*Candidatus* Nitrotoga", *Lentisphaera, Nitrosospira, Modestobacter, Conexibacter, Dokdonella, Gryllotalpicola, Granulicella, Actinospica, Kouleothrix, Aquisphaera, Glycomyces,* "*Candidatus* Amphibiichlamydia", "*Candidatus* Xiphinematobacter", and *Mucilaginibacter*. *Acidibacter, Opitutus, Mycobacterium, Phenylobacterium, Reyranella, Dongia, Pedomicrobium, Steroidobacter, Alterococcus, Roseimicrobium, Byssovorax,* "*Candidatus* Alysiosphaera", AKYG587 of *Phycisphaeraceae, Ramlibacter, Gimesia, Methylobacter, Desulfonatronum, Hymenobacter, Acinetobacter, Tuwongella, Phormidium* MBIC10003, *Adhaeribacter, Labilithrix, Anaerovorax, Solitalea, Actinomadura, Cephaloticoccus,* wb1-P19 of *Nitrosococcaceae, Desulfosporosinus, Egibacter,* ADurb.Bin063-1 of *Pedosphaeraceae, Microvirga, Pontibacter, Chryseobacterium, Herbaspirillum, Tenggerimyces, Fervidobacterium, Kallotenue, Lachnoclostridium, Geotalea, Aureispira, Desulfurispora,* and *Sideroxydans* were enriched in acidic substrate only in the bulk soil habitat. These taxa are likely more adapted to acidic condition than to alkaline.


A)
Bulk soil
Alkaline
Acidic
Rhizosphere
Alkaline
Acidic
Roots
Alkaline
Acidic

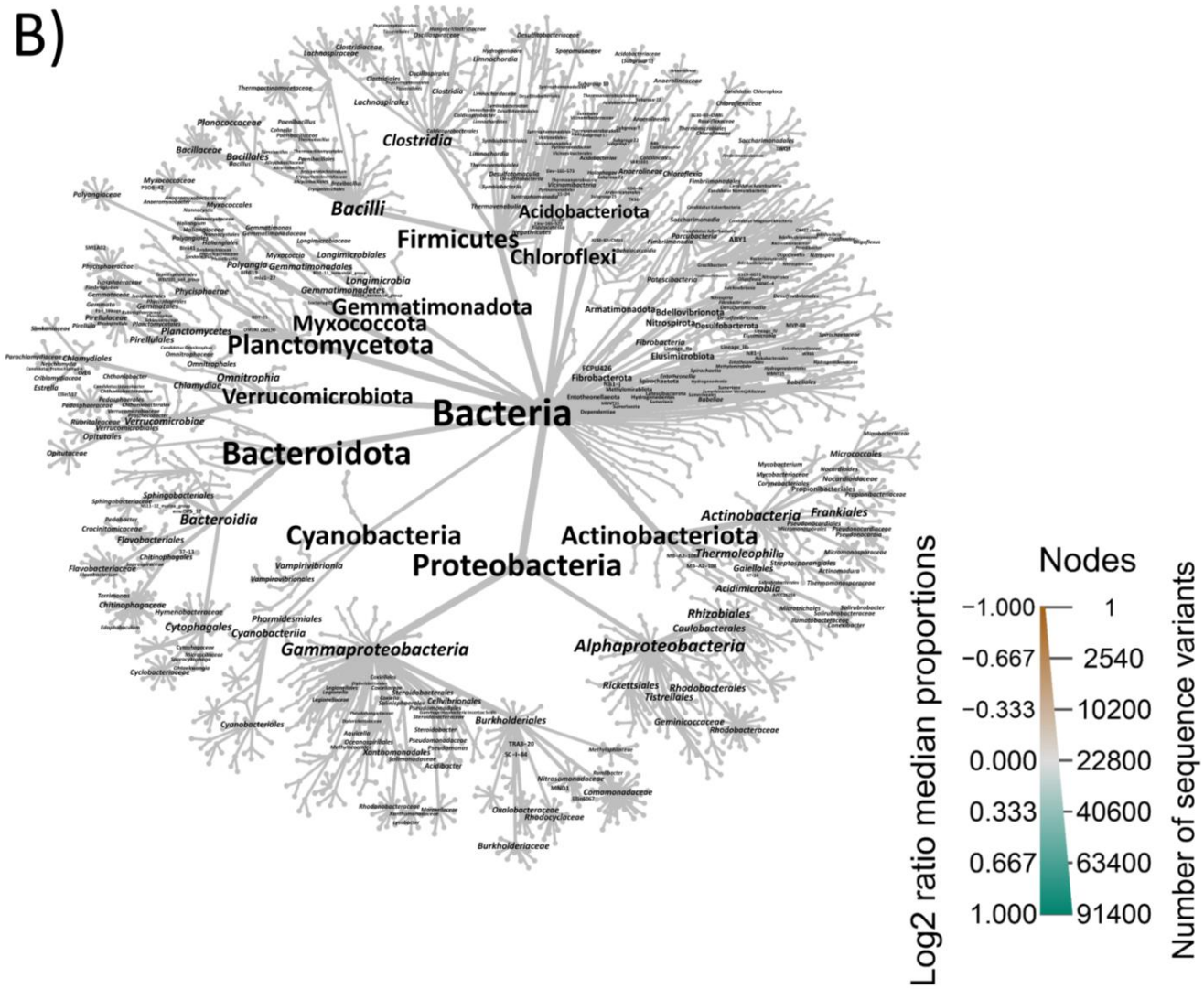


Figure 32 – Pairwise comparison of bacterial communities between alkaline and acidic substrate of bulk soil, rhizosphere, and root samples in the controlled greenhouse experiment, from phylum up to genus level (A). Wilcoxon rank-sum test, followed by a Benjamini-Hochberg (FDR) correction for multiple comparisons, were performed to determine significant differences ($p < 0.05$). Significantly enriched taxa were colored according to the soil substrate (acidic or alkaline) where they were found to be enriched in and the color intensity reflects log-2 ratio in median proportions (the values are depicted in Figure B). Figure B depicts the key tree with the taxonomic information. Node diameter reflects number of sequence variant classified as that taxon whereas edge width reflects number of reads.

Sporichthya, Rubrivivax, I-8 of *Phycisphaeraceae, Asticcacaulis, Neochlamydia, Thermomonas,* UTBCD1 of *Chitinophagaceae, "Candidatus* Mesochlamydia"*, Crocinitomix, Duganella, Oscillochloris, Telmatospirillum, Desulfovibrio, Minicystis, Herpetosiphon, Thiomonas,* and *Endozoicomonas* were enriched in acidic substrate only in the rhizosphere. Since they were not alkaline-enriched in bulk soil habitat, their enrichment may also be related to the host plant species.

Furthermore, *Rhizobacter, Hirschia, Flavihumibacter,* MND1 and IS-44 of *Nitrosomonadaceae, Devosia, Pirellula,* OLB14 of *Anaerolineales, Telmatocola, Paenibacillus, Variovorax, Oscillatoria* PCC-6304, *Chitinimonas,* OLB12 of *Microscillaceae, Lautropia, Xenophilus, Burkholderia-Caballeronia-Paraburkholderia* group, YC-ZSS-LKJ63 of *Hydrogenedensaceae, Deferrisoma, Paludibaculum, Marinobacter, Sumerlaea, Tagaea, Brevundimonas,* MM2 of *Methylophilaceae, Leptospirillum, Halobacillus, Salinicola, Reichenbachiella, Nocardiopsis,* and *Thiovirga* were enriched in acidic substrate but only inside the roots. Again, this enrichment inside the roots may be due to other selective factors such as nutrient availability or bacterial adaptation, as the pH inside the roots is most likely constant.

To investigate the impact of plant species on bacterial composition in the rhizosphere, comparison of rhizosphere bacterial communities between different plant species (*H. perforatum, H. olympicum,* and *H. balearicum*) at phylum level (except for the two abundant classes *Alphaproteobacteria* and *Gammaproteobacteria*) was conducted (Figure 33). The result revealed that Actinobacteriota and Firmicutes were significantly depleted for *H. balearicum* but only in the total bacterial communities (DNA-based), when compared to the other two plant species (Figure 33). This implies that although Firmicutes were not that abundant (DNA-based) in the rhizosphere of *H. balearicum*, their metabolic activity could be higher compared to other phyla/classes. Thus, their abundance in the active (RNA-based) bacterial communities in the rhizosphere of *H. balearicum* was not significantly different when compared to other plant species. None of the phlya/classes were significantly enriched/depleted for a particular plant species when compared to both remaining plant species in the active (RNA-based) bacterial communities, implying the necessity to observe the bacterial composition at a deeper taxonomy level.

In order to determine the bacterial taxa (from phylum up to genus level) that were enriched in the rhizosphere of a particular plant species when individually compared with other plant species, pairwise comparison was conducted for each possible pairwise combination of the three plant species (Figure 34). The host plant specificity may emphasize the importance of the bacteria to the specific ecological roles carried out by the host plant.

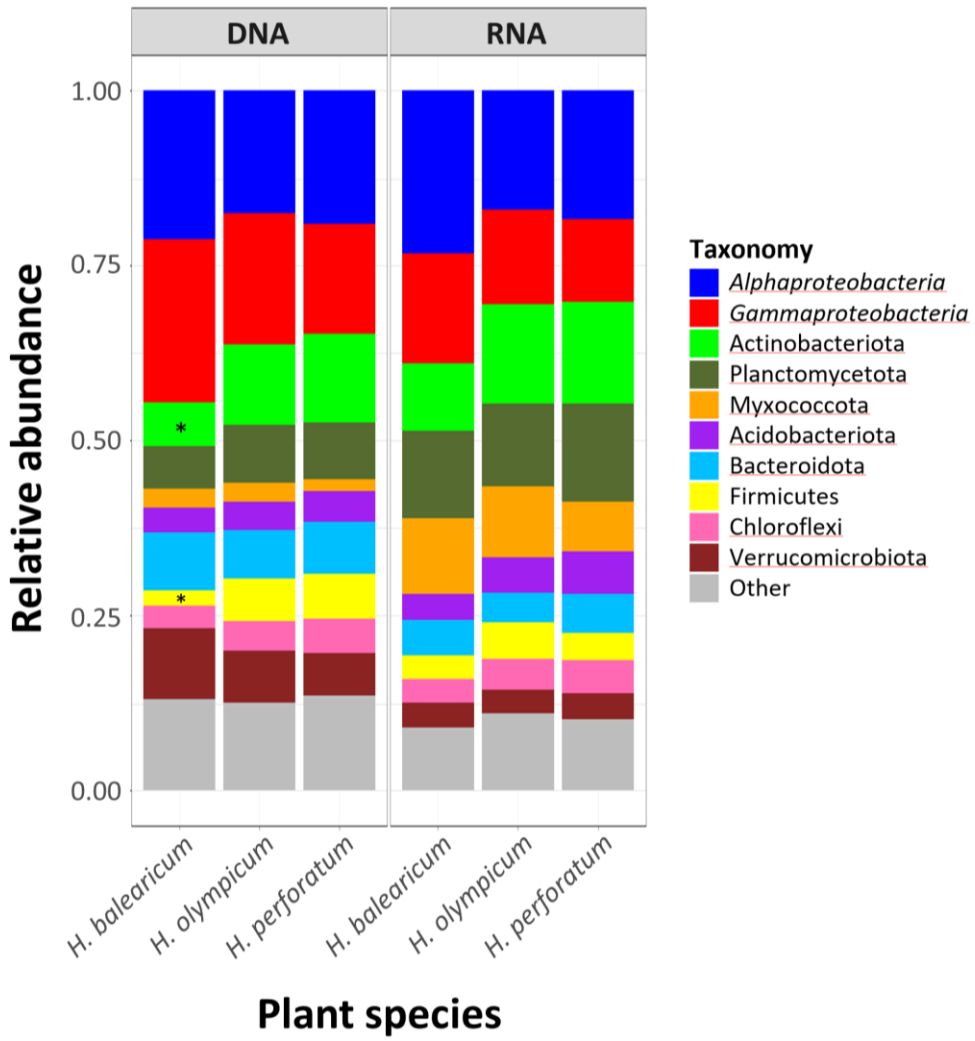


Figure 33 – Relative abundance of bacterial phyla and *Alphaproteobacteria* and *Gammaproteobacteria* in the rhizosphere of *Hypericum* plants from the controlled greenhouse experiment. Significantly enriched/depleted taxa for a particular plant species, when compared to the other two, are marked with asterisks ($p < 0.05$; multcomp test). Both total (DNA-based) and active (RNA-based) bacterial communities were investigated.

More taxa were differentially enriched in the total (DNA-based) bacterial communities, while less differences could be observed in the active (RNA-based) communities (as implied with less colored nodes and edges on the pairwise-trees of the active communities). The result implies that although many rhizosphere taxa were differentially abundant between distinct plant species, these only represent dormant taxa or extracellular DNA since the pattern was not observed in the active (RNA-based) bacterial communities. *H. perforatum* and *H. olympicum* rhizosphere shared a higher degree of similarity of bacterial communities than when each of the plant species was individually compared with *H. balearicum* (fewer colored nodes and edges on the pairwise-trees). Since both

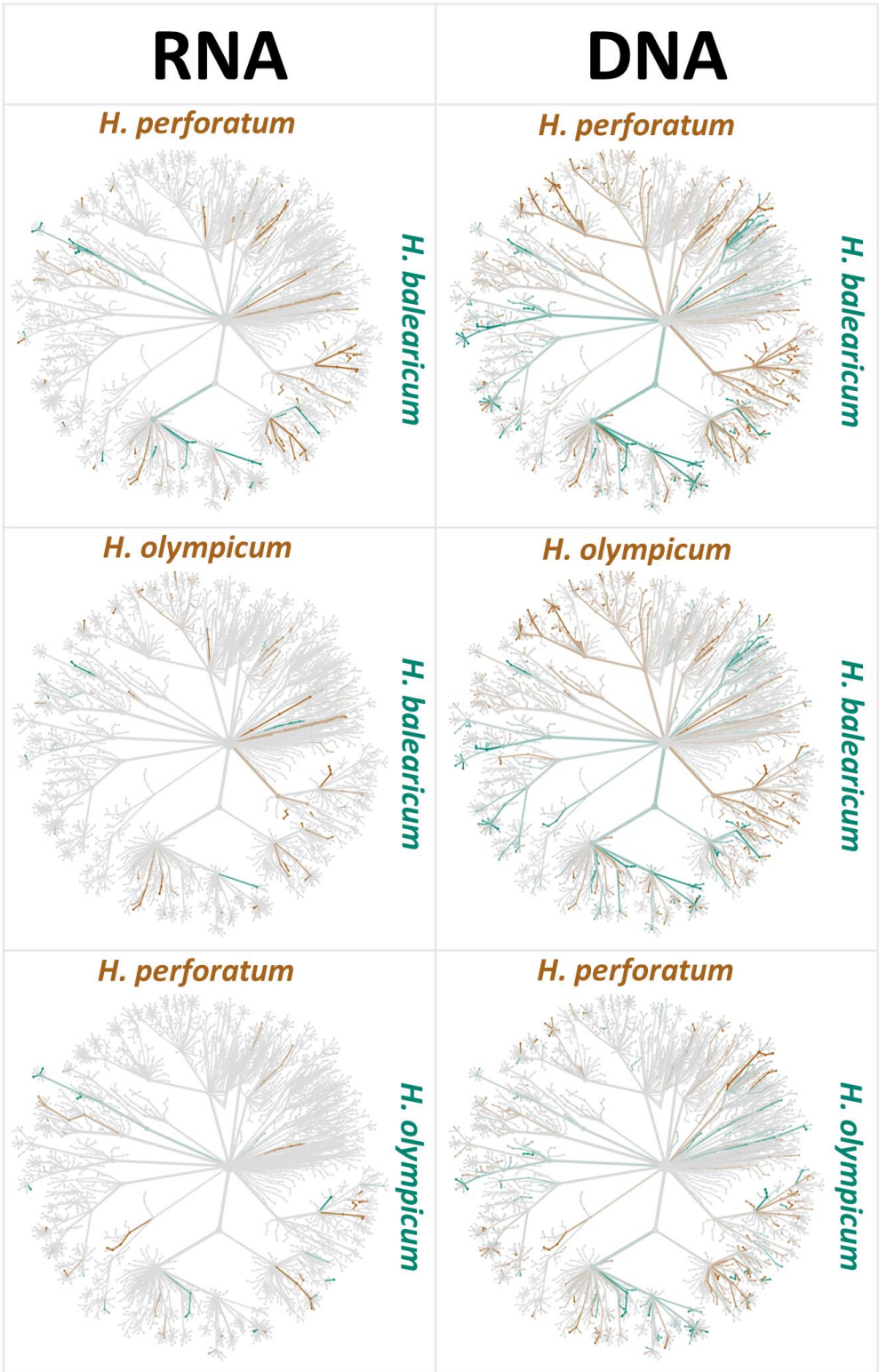
plant species are more closely related to each other in comparison to *H. balearicum,* they may select similar subset of bacterial taxa in their rhizosphere.

Fibrobacterota, Entotheonellaeota, and Methylomirabilota phyla along with genera such as *Mycobacterium*, *Pseudonocardia*, *Georgfuchsia*, "*Candidatus* Entotheonella", *Laceyella, Methylocaldum, Fictibacillus,* mle1-7 of *Nitrosomonadaceae,* Pir3 lineage of *Pirellulaceae, Methylobacter, Thiorhodovibrio, Thermincola, Sedimentibacter,* and *Thermoflavimicrobium* were among the taxa enriched in the rhizosphere of *H. perforatum* and *H. olympicum* when compared with *H. balearicum*, in both total (DNA-based) and active (RNA-based) bacterial communities. Some other taxa that were enriched in the rhizosphere of the two plant species but only in RNA data included *Pancagrimonas*, *Desulfonatronum*, *Adhaeribacter, Elioraea, Solitalea, Azohydromonas, Actibacterium, Eubacterium, Oceanobacillus* and *Desulfosporosinus*. Since *H. perforatum* and *H. olympicum* produce hypericin and hyperforin, these taxa may have the ability to modulate the secondary metabolites of the host plant species. However, this result needs to be validated in the future.

On the other hand, *Methylotenera, Cellvibrio, Luteolibacter, Flavihumibacter,* DSSD61 of *Nitrosomonadaceae, Coxiella, Pseudoflavitalea,* and *Imperialibacter* were consistently enriched in the rhizosphere of *H. balearicum* in the total (DNA-based) bacterial communities, and *Lacunisphaera* and *Pseudohongiella* in the active (RNA-based) bacterial communities when compared with each of the two other plant species. These taxa may perform specific functions related to the ecological roles of *H. balearicum* that are not including hypericin and hyperforin production.

Pseudonocardia, *Arenibacter, Nakamurella, Falsirhodobacter, Burkholderia-Caballeronia-Paraburkholderia* group, *Rubellimicrobium, Edophabaculum, Chiayiivirga,* and *Fulvivirga* were the genera that were enriched in the rhizosphere of *H. perforatum* when compared to *H. olympicum* in both total (DNA-based) and active (RNA-based) bacterial communities whereas *Aeromicrobium, Ahniella, Microlunatus, Bryobacter, Xenophilus,* CL500-3 of *Phycisphaeraceae, Gimesia, Algisphaera, Tuwongella,* and *Fluviicola* were enriched in the rhizosphere of *H. perforatum* only in the active bacterial communities (RNA-based). *Phenylobacterium, Pedomicrobium, Nocardioides, Ilumatobacter, Georgfuchsia, Terrimicrobium, Rhodopirellula, Aquicella, Thalassobaculum, Taibaiella, Puia, Hymenobacter, Rhodocytophaga, Owenweeksia, Subsaxibacter, Microbacterium,* and "*Candidatus* Amoebophilus" were enriched for *H. perforatum* only in the total bacterial communities (DNA-based), implying that these taxa are likely dormant or represent relic DNA.

A)


RNA
DNA
H. perforatum
H. balearicum
H. perforatum
H. balearicum
H. olympicum
H. balearicum
H. olympicum
H. balearicum
H. perforatum
H. olympicum
H. perforatum
H. olympicum

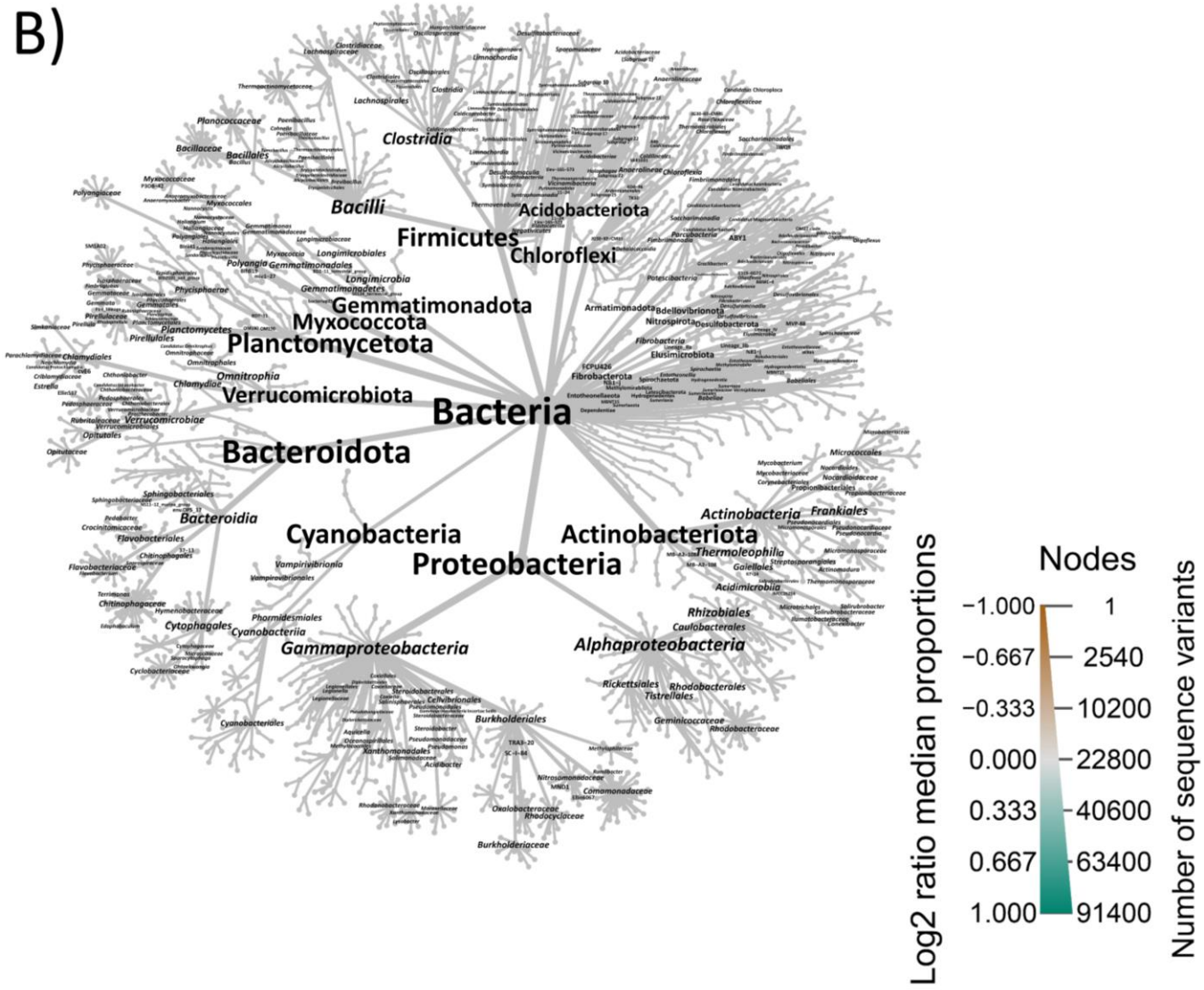


Figure 34 – Pairwise comparison of total (DNA-based) and active (RNA-based) rhizosphere bacterial communities between different plant species in the controlled greenhouse experiment, from phylum up to genus level (A). Wilcoxon rank-sum test, followed by a Benjamini-Hochberg (FDR) correction for multiple comparisons, were performed to determine significant differences ($p < 0.05$). Significantly enriched taxa were colored according to the plant species where they were found to be enriched in and the color intensity reflects log-2 ratio in median proportions (the values are depicted in Figure B). Figure B depicts the key tree with the taxonomic information. Node diameter reflects number of sequence variant classified as that taxon whereas edge width reflects number of reads.

On the other hand, the enriched taxa in the rhizosphere of *H. olympicum* when compared to *H. perforatum* in both total (DNA-based) and active (RNA-based) bacterial communities included *Rhizobacter*, *Acidibacter*, *Opitutus*, *Sphingobium,* P3OB-42 of *Myxococcaceae, Byssovorax, Chryseolinea,* and *Massilia*. In addition, *Polyangium, Tychonema* CCAP 1459-11B, and *Paracoccus* were the *H. olympicum*-enriched taxa only in the active bacterial communities (RNA-based) whereas IMCC26134 of *Opitutaceae*, *Leptospira,* and *Nitrosomonas* were the enriched taxa only in the total communities (DNA-based). The taxa that were enriched in the rhizosphere of *H. olympicum* in the active bacterial communities (RNA-based) or in both communities may contribute to specific ecological roles carried out by *H. olympicum* but not by *H. perforatum.*

Aside from the rhizosphere, the impact of plant species on bacterial composition was also determined for root habitat at phylum/class level (Figure 35). In the total (DNA-based) bacterial communities, Acidobacteriota and Chloroflexi were significantly enriched in the roots of *H. perforatum* when compared to *H. olympicum* and *H. balearicum*. Since no significant differences could be observed in the active communities (RNA-based), these phyla are not necessarily more active for *H. perforatum*. Moreover, Firmicutes was depleted in the roots of *H. balearicum* when compared to both remaining plant species in the active bacterial communities (RNA-based), impying that the phylum is particularly inactive inside the roots of *H. balearicum*.

In order to evaluate further the impact of plant species on the total (DNA-based) and active (RNA-based) bacterial communities inside the roots, pairwise comparison of the root (endophytic) bacterial communities, up to genus level, was conducted for each possible pairwise combination of the three plant species (Figure 36). Number of taxa enriched between distinct plant species was not as prominent for root bacterial communities, as had been observed for the rhizosphere communities, even though the effect of plant species on both habitats was similar based on variance partitioning (approximately 20%; Figure 27). This may be explained by lower bacterial diversity inside the roots when compared to the rhizosphere, as confirmed with the result of alpha diversity (Figure 24).

Hyphomicrobium, *Mycobacterium*, *Lysobacter, Georgfuchsia,* and *Aquicella* were enriched in the roots of *H. perforatum* and *H. olympicum* when each of the two plant species were individually compared with *H. balearicum,* in both total (DNA-based) and active (RNA-based) bacterial communities, whereas *Lysobacter* was enriched in the roots of the two plant species specifically in the active (RNA-based) communities. These taxa may be affiliated with the ecological roles performed by both plant species, including elicitation of hypericin and hyperforin, but this needs to

be validated further. Moreover, *Gemmata* was enriched for the two plant species but only in the total bacterial communities (DNA-based) and thus may imply dormancy.

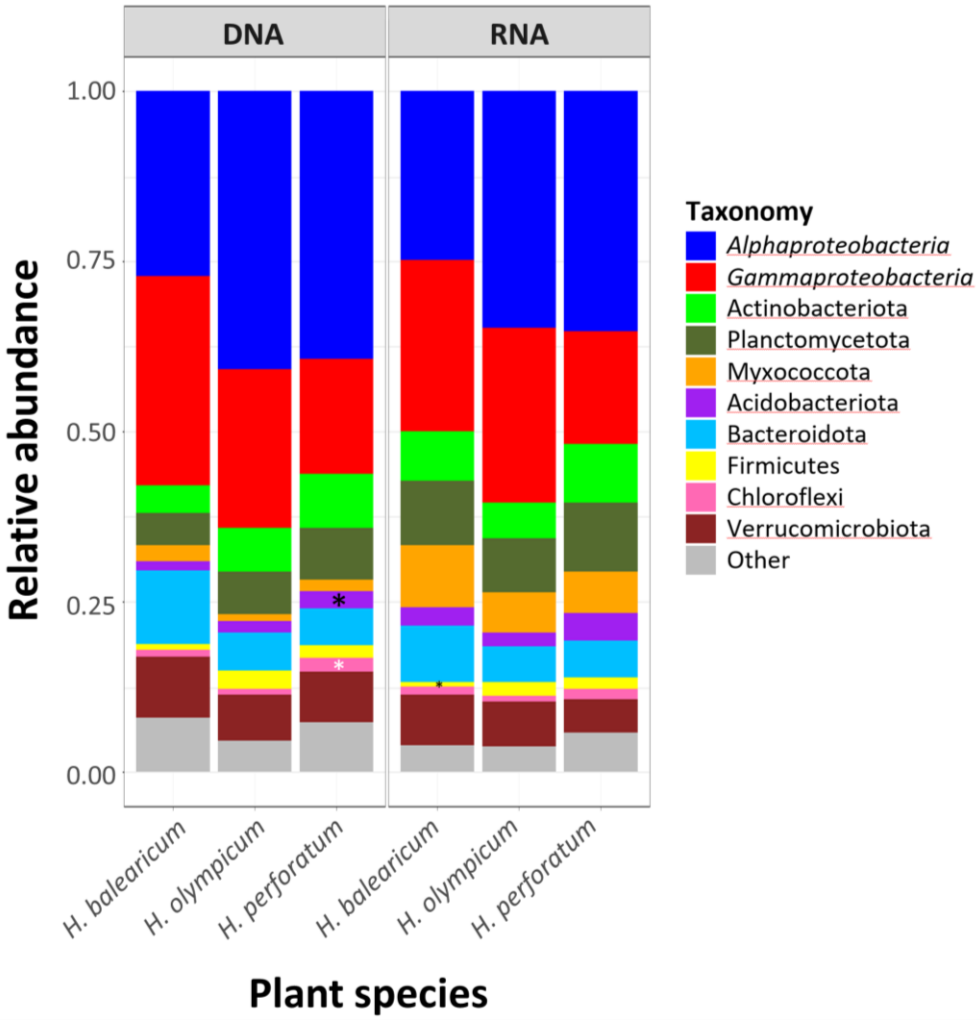


Figure 35 – Relative abundance of bacterial phyla and *Alphaproteobacteria* and *Gammaproteobacteria* inside the roots of *Hypericum* plants from the controlled greenhouse experiment. Significantly enriched/depleted taxa for a particular plant species, when compared to the other two, are marked with asterisks ($p < 0.05$; multcomp test). Both total (DNA-based) and active (RNA-based) bacterial communities were investigated.

Sphingomonas, *Phenylobacterium*, *Pedomicrobium, Pseudoxanthomonas, Pseudonocardia, Lysinibacillus, Methylibium,* "*Candidatus* Entotheonella", *Arenimonas, Nakamurella,* "*Candidatus* Alysiosphaera", *Falsirhodobacter, Methylocaldum, Methylobacter,* Pir2 lineage of *Pirellulaceae*, *Hymenobacter, Legionella, Chiayiivirga,* and *Parvibaculum* were the enriched taxa in the roots of *H. perforatum* when compared to *H. balearicum* in both total (DNA-based) and active (RNA-based) bacterial communities, whereas *Bacillus, Nocardioides, Ureibacillus, Dyadobacter,*

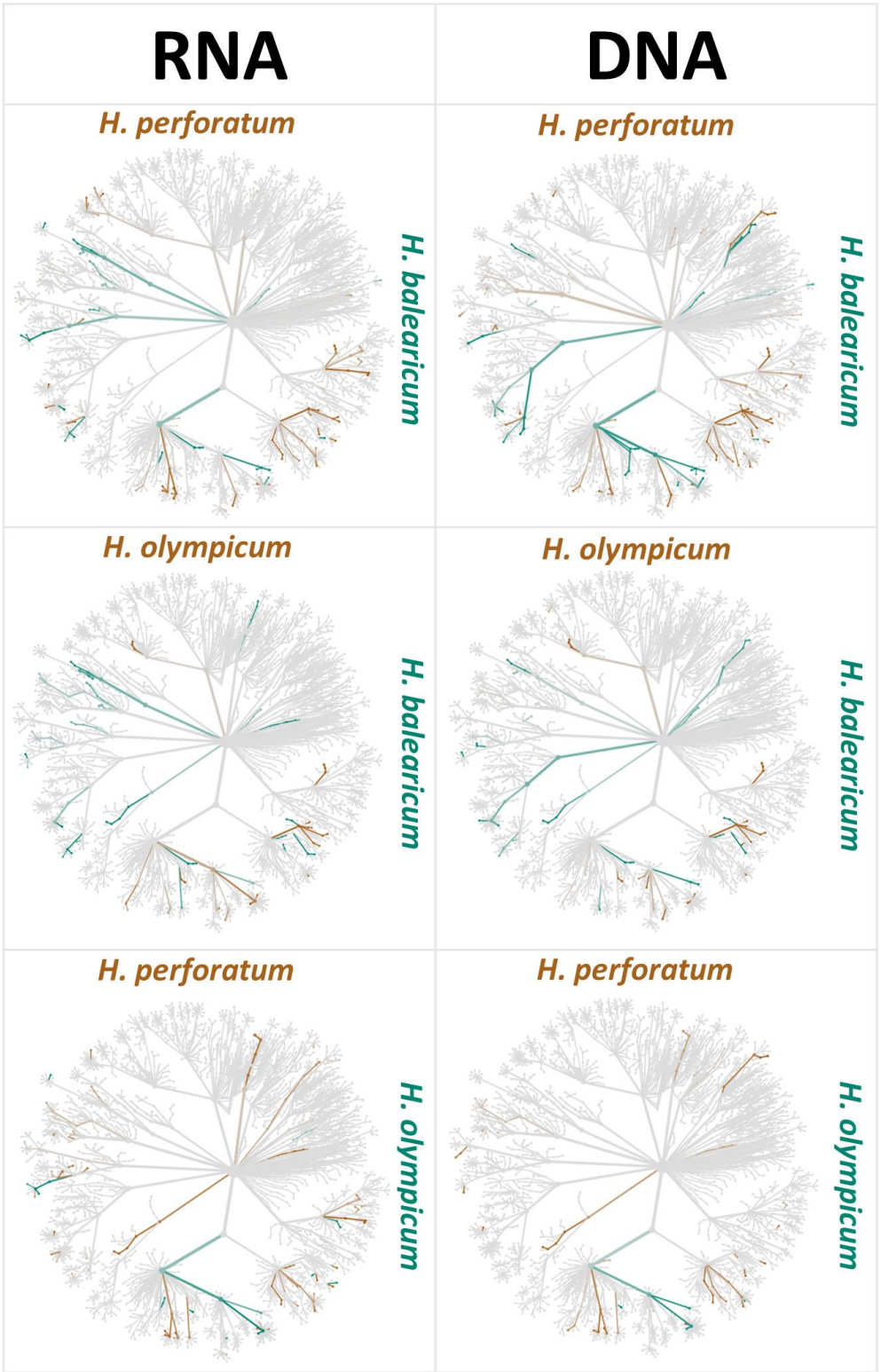
Paenisporosarcina, Tahibacter, Anaeromyxobacter, Arenibacter, Lautropia, Ramlibacter, Fictibacillus, Brevibacillus, Aneurinibacillus, Cohnella, Hylemonella, Parachlamydia, Neochlamydia, Taibaiella, Ruminiclostridium, Fulvivirga, and *Amaricoccus* were specifically enriched in the active communities (RNA-based). In addition, *Dongia, Roseimicrobium, Pajaroellobacter, Estrella, Rhodobacter, Rhodopirellula, Thermopolyspora,* AKYG587 of *Phycisphaeraceae,* IMCC26207 of *Microtrichaceae, Blastopirellula, Sphaerobacter, Thermobifida, Actinomadura, Simkania,* and *Defluviimonas* were enriched inside the roots of *H. perforatum* but only in total (DNA-based) bacterial communities and thus may represent dormant cells or extracellular DNA.

Rhizobacter, *Allorhizobium-Neorhizobium-Pararhizobium-Rhizobium* group, *Pseudomonas, Acidovorax, Erysipelatoclostridium,* and *Gemmobacter* were specifically enriched inside the roots *H. olympicum*, when compared to *H. balearicum,* in both total (DNA-based) and active (RNA-based) bacterial communities whereas *Pseudorhodoplanes, Bosea, Massilia,* and *Pantoea* were enriched only in the active bacterial communities (RNA-based). In addition, *Chitinimonas, Vogesella,* and *Asticcacaulis* were enriched inside the roots of *H. olympicum* but only in the total (DNA-based) bacterial communities and thus may be inactive.

On the other hand, genera including *Cellvibrio*, *Niastella*, and *Imperialibacter* were consistently enriched inside the roots of *H. balearicum* when compared to both remaining plant species, in both total (DNA-based) and active (RNA-based) bacterial communities whereas SWB02 of *Hyphomonadaceae*, *Lacunisphaera,* and *Marinoscillum* were specifically enriched in the active communities (RNA-based). Thus, these bacterial taxa may be affiliated with the ecological roles of *H. balearicum* (excluding hypericin and hyperforin, as the plant species is a non-producer).

Between *H. perforatum* and *H. olympicum*, *Phenylobacterium*, *Pedomicrobium, Pseudoxanthomonas,* SM1A02 of *Phycisphaeraceae, Methylibium, Arenibacter, Nakamurella,* "*Candidatus* Alysiosphaera", *Falsirhodobacter, Rhodopirellula, Ereboglobus,* and *Chiayiivirga* were enriched inside the roots of *H. perforatum,* in both total (DNA-based) and active (RNA-based) bacterial communities, whereas *Sphingomonas, Reyranella, Nocardioides, Pseudonocardia, Planctopirus, Nannocystis, Terrimicrobium, Tahibacter, Microlunatus, Arenimonas, Lautropia, Bdellovibrio, Ramlibacter, Fictibacillus, Edaphobaculum, Hylemonella, Parachlamydia, Gimesia, Taibaiella, Hymenobacter, Legionella,* "*Candidatus* Fritschea"*, Oscillochloris, Parvibaculum, Fulvivirga,* and *Amaricoccus* were enriched specifically in the active (RNA-based) bacterial communities. In addition, *Pajaroellobacter, Estrella, Thermopolyspora, Methylocaldum,* AKYG587 of *Phycisphaeraceae, Thalassobaculum, Glycomyces,* and *Defluviimonas* were enriched inside the roots of *H. perforatum* but only in the total (DNA-based) bacterial communities and thus may be inactive.

A)


RNA
DNA
H. perforatum
H. balearicum
H. perforatum
H. balearicum
H. olympicum
H. balearicum
H. olympicum
H. balearicum
H. perforatum
H. olympicum
H. perforatum
H. olympicum

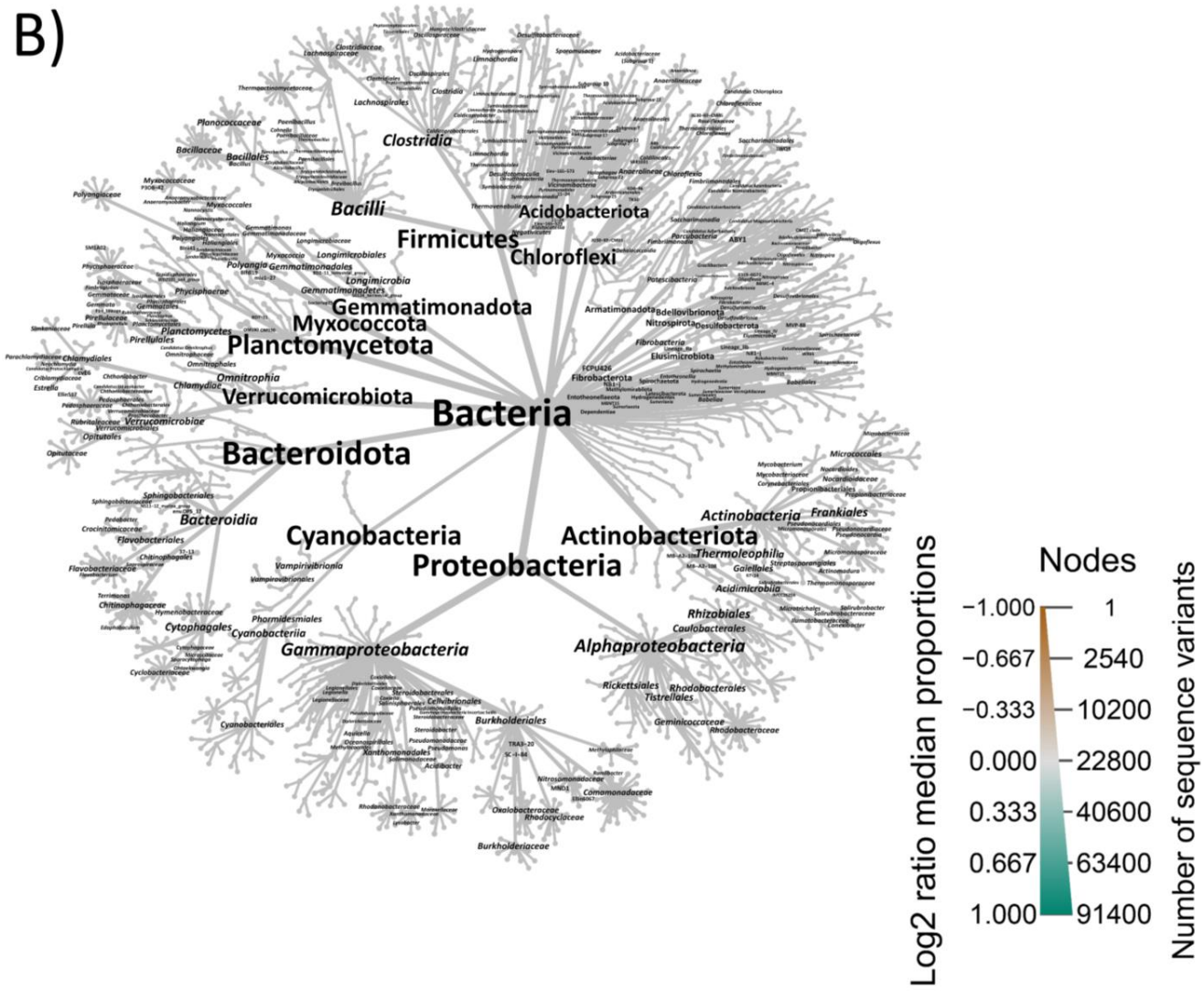


Figure 36 – Pairwise comparison of total (DNA-based) and active (RNA-based) root (endophytic) bacterial communities between different plant species in the controlled greenhouse experiment, from phylum up to genus level (A). Wilcoxon rank-sum test, followed by a Benjamini-Hochberg (FDR) correction for multiple comparisons, were performed to determine significant differences ($p < 0.05$). Significantly enriched taxa were colored according to the plant species where they were found to be enriched in and the color intensity reflects log-2 ratio in median proportions (the values are depicted in Figure B). Figure B depicts the key tree with the taxonomic information. Node diameter reflects number of sequence variant classified as that taxon whereas edge width reflects number of reads.

On the other hand, *Rhizobacter* and *Pseudomonas* were enriched inside the roots of *H. olympicum,* in both total (DNA-based) and active (RNA-based) bacterial communities, whereas *Opitutus*, IMCC26134 of *Opitutaceae, Sphingobium, Ketobacter, Byssovorax, Massilia,* and *Asticcacaulis* were specifically enriched in the active (RNA-based) communities.

Based on the comparison between total (DNA-based) and active (RNA-based) bacterial communities that has been conducted earlier, we could try to predict which taxa are active in each habitat. However, we only performed the analysis up to genus level and thus not represent an individual bacterium. Due to the fact that the metabolic capability is varied across the same genus, we employed a specific threshold based on the rRNA:rDNA ratio to identify transcriptionally active taxa at the deepest taxonomic level possible (sequence variant level; see sub-chapter 3.1.7.3 and Figure 17 for detail).

Of the active sequence variants with rRNA:rDNA ratio values above the threshold, a low number were detected specifically in the active (RNA-based) communities while the majority were detected in both total (DNA-based) and active (RNA-based) bacterial communities (Figure 37A). The result implies that the DNA-based community analysis is able to detect potentially active taxa. However, the DNA abundance does not reflect activity and thus activity-based analysis needs to be employed to assist the DNA-based community study. Moreover, although root (endophytic) bacterial communities were less diverse when compared to the other two habitats (Figure 24), the percentage of active sequence variants inside the root was the highest, especially on alkaline substrate (Figure 37B). The higher number of active sequence variants inside the roots may be related to the plant dependency on the endophytic bacteria to maintain its health and productivity.

The group-active sequence variants were identified for each group of bulk soil, rhizosphere, and inside roots of a particular plant species, as the ones that were active in at least 2 biological replicates (see subchapter 3.1.7.4 and Figure 19 for detail). In total, there were 824 sequence variants that fitted the criterium. Each group contained phylogenetically diverse active taxa (Figure 38). Members of *Streptomyces,* Blrii41 of *Myxococcota, Polyangium,* and *Haliangium* were among the group-active taxa in all bulk soil groups regardless of soil substrate and plant species. These taxa are potentially important contributors of soil ecosystem functioning. Moreover, sequence variants belonging to OM27 clade of Bdellovibrionota, *Telmatocola,* TK10 of Chloroflexi, *Sphingomonas,* and *Dactylosporangium* were among the bulk soil active taxa only on acidic substrate, independently of plant species, whereas members of *Pseudomonas, Methylibium, Fimbriiglobus,* and Euzebya were active only on alkaline substrate. These results emphasize different preferences of bacteria on distinct soil substrate with different pH.

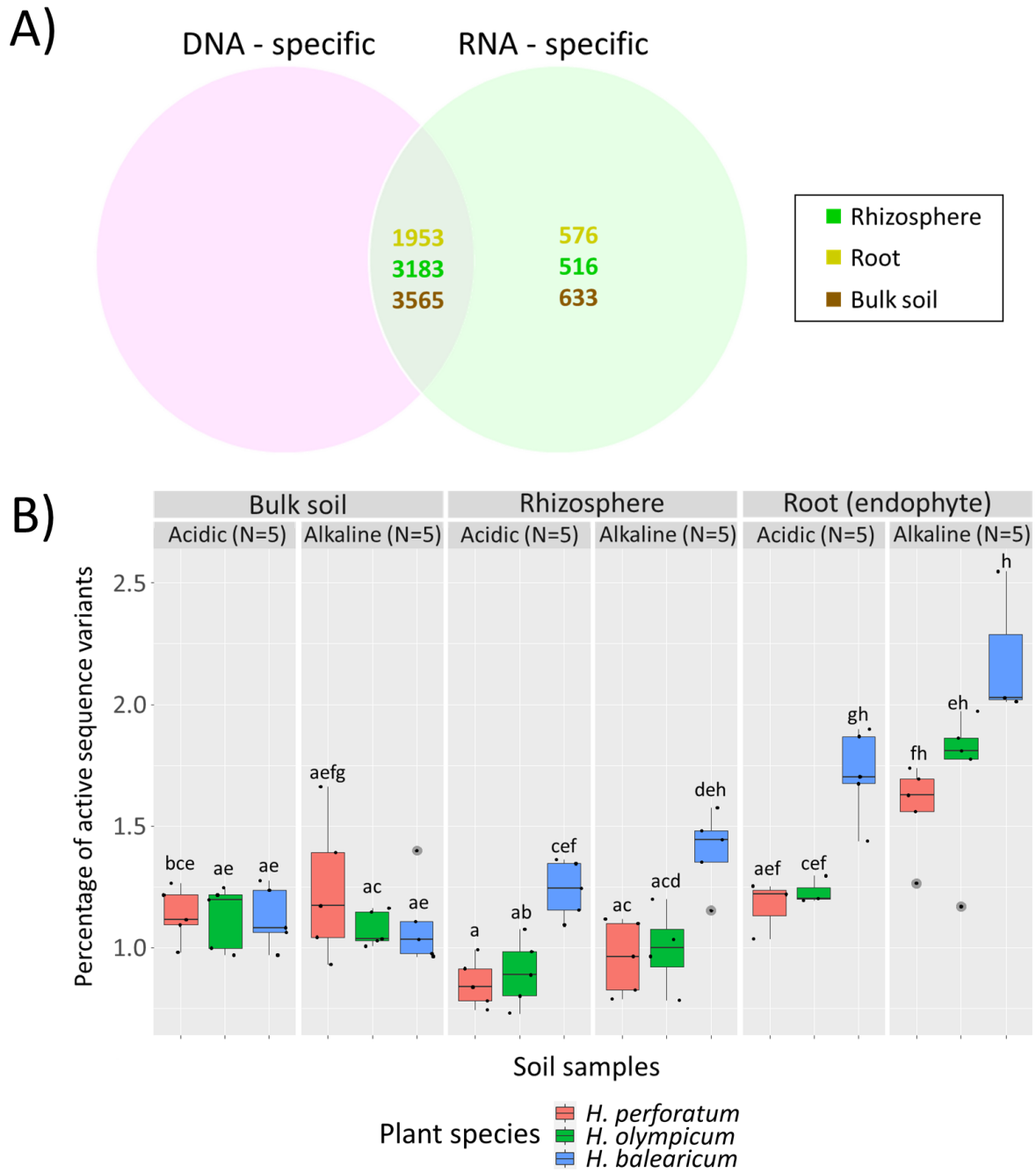


Figure 37 – The counts of active sequence variants (with rRNA:rDNA value above the threshold) that were detected only in the active (RNA-based) bacterial communities (A; on the right circle) together with the ones that were detected in both total (DNA-based) and active (RNA-based) bacterial communities (A; intersection of venn diagram) (A). None of sequence variants that present only in the total communities passed the threshold (rRNA:rDNA ratio value = 0). B) The percentage of active sequence variants in bulk soil, the rhizosphere, and inside roots of *H. perforatum, H. olympicum,* and *H. balearicum* from the controlled greenhouse experiment. Grey background depicts an outlier. Letters denote significant differences between different habitat and soil substrate (multcomp, $p < 0.05$). N represents the number of samples for each group.

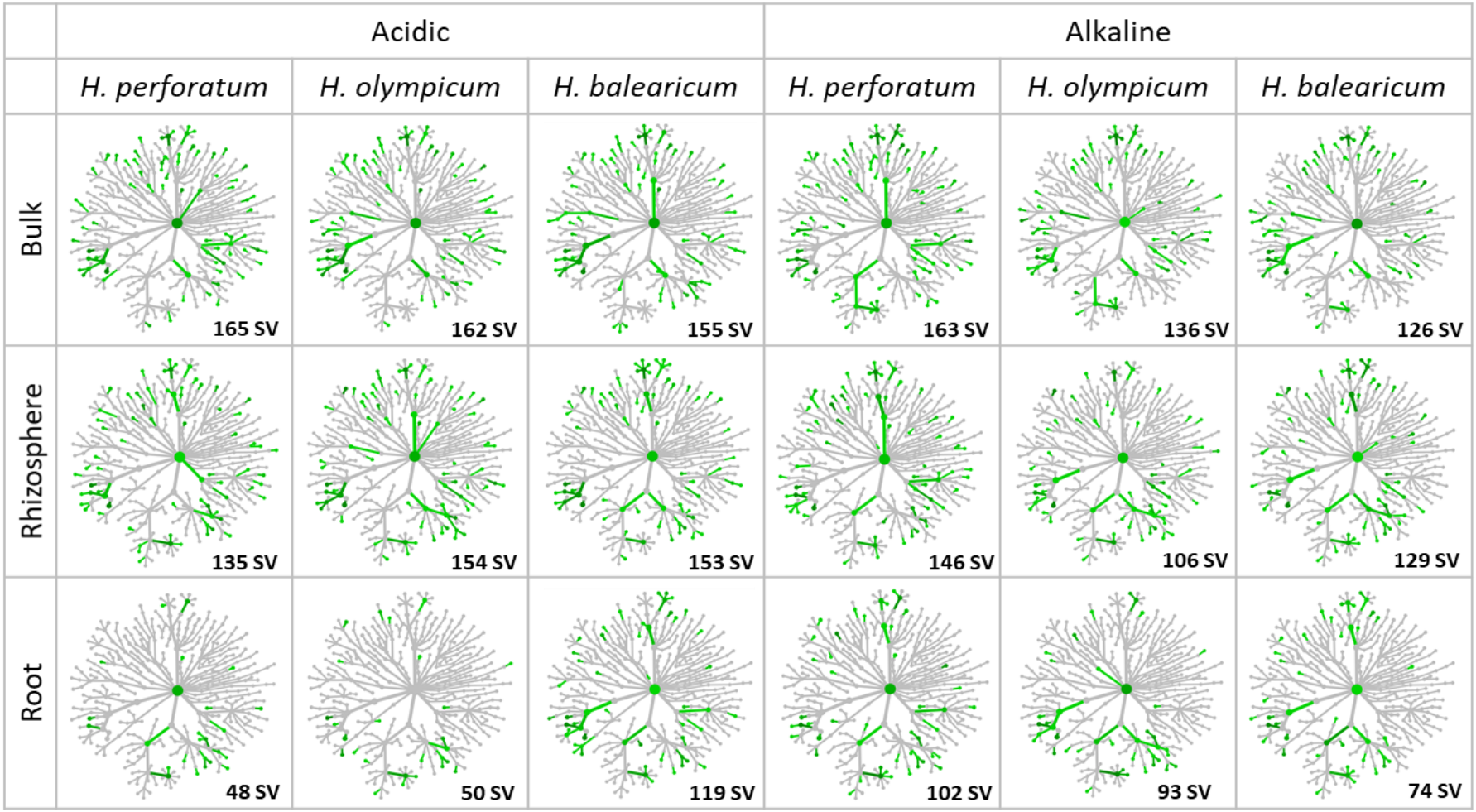
The rhizosphere active taxa included members of *Streptomyces,* PMMR1 of *Caulobacteriaceae, Phenylobacterium,* and Blrii41, independently of soil substrate and plant species. In addition, sequence variants of *Telmatocola* and TK10 of Chloroflexi were active only on acidic substrate and sequence variants of *Pseudorhodoplanes* and *Sporocytophaga* were active only on alkaline substrate. These taxa are likely important for plants in general since they were active in the rhizosphere of all plant species. Soil substrate specificity may reflect the preferences of this taxa growing at specific pH range.

Of root samples, sequence variants of PMMR1 of *Caulobacteriaceae, Phenylobacterium, Ohtaekwangia* and *Haliangium* were consistently active across all root groups, independently of soil substrate and plant species, implying their importance to host plants in general. In addition, members of *Pelomonas* was specifically active on alkaline substrate while no taxon was observed to be acidic-specific. Since the pH inside the roots is likely constant, other selective factors may drive *Pelomonas* to be active inside the roots when plants grew on alkaline substrate (may be related to the response of the host plant growing on alkaline condition).

The sequence variants that were specifically active in the rhizosphere of hypericin and hyperforin-producing plant species (*H. perforatum* and *H. olympicum*) and a non-producer (*H. balearicum*) were investigated on both acidic and alkaline substrates. The criteria to determine this active-specific taxa include: i selection based on the threshold, ii being active at least in 2 biological replicates, iii being active in the rhizosphere of producer but not in the rhizosphere of non-producer, and iv DNA needs to be detected in all bulk soil samples. This implies the importance of these taxa to their corresponding host plant, which may be related to the production of hypericin and hyperforin in *H. perforatum* and *H. olympicum.*

The rhizosphere active taxa that were specific for hypericin and hyperforin-producing species included members of *Bacillus* and *Pedomicrobium* on acidic substrate. Many taxa could not be identified at genus level including members of P3OB-42 of *Myxococcaceae, Gemmataceae, Entotheonellaceae, Blrii41* of *Polyangiales*, *Rhizobiales, Thermovenabulales*, SBR1031 of Chloroflexi and PAUC26f of Acidobacteriota (Figure 39). On alkaline substrate, the active taxa included members of *Solirubrobacter*, Blrii41 of *Polyangiales*, IMCC26256 of *Acidimicrobiia*, *Microtrichales*, and *Frankiales*. These taxa may elicit hypericin and hyperforin in *H. perforatum* and *H. olympicum* but this result needs to be validated in the future.

A)


Acidic
Alkaline
H. perforatum
H. olympicum
H. balearicum
H. perforatum
H. olympicum
H. balearicum
Bulk
Rhizosphere
Root
165 SV
162 SV
155 SV
163 SV
136 SV
126 SV
135 SV
154 SV
153 SV
146 SV
106 SV
129 SV
48 SV
50 SV
119 SV
102 SV
93 SV
74 SV

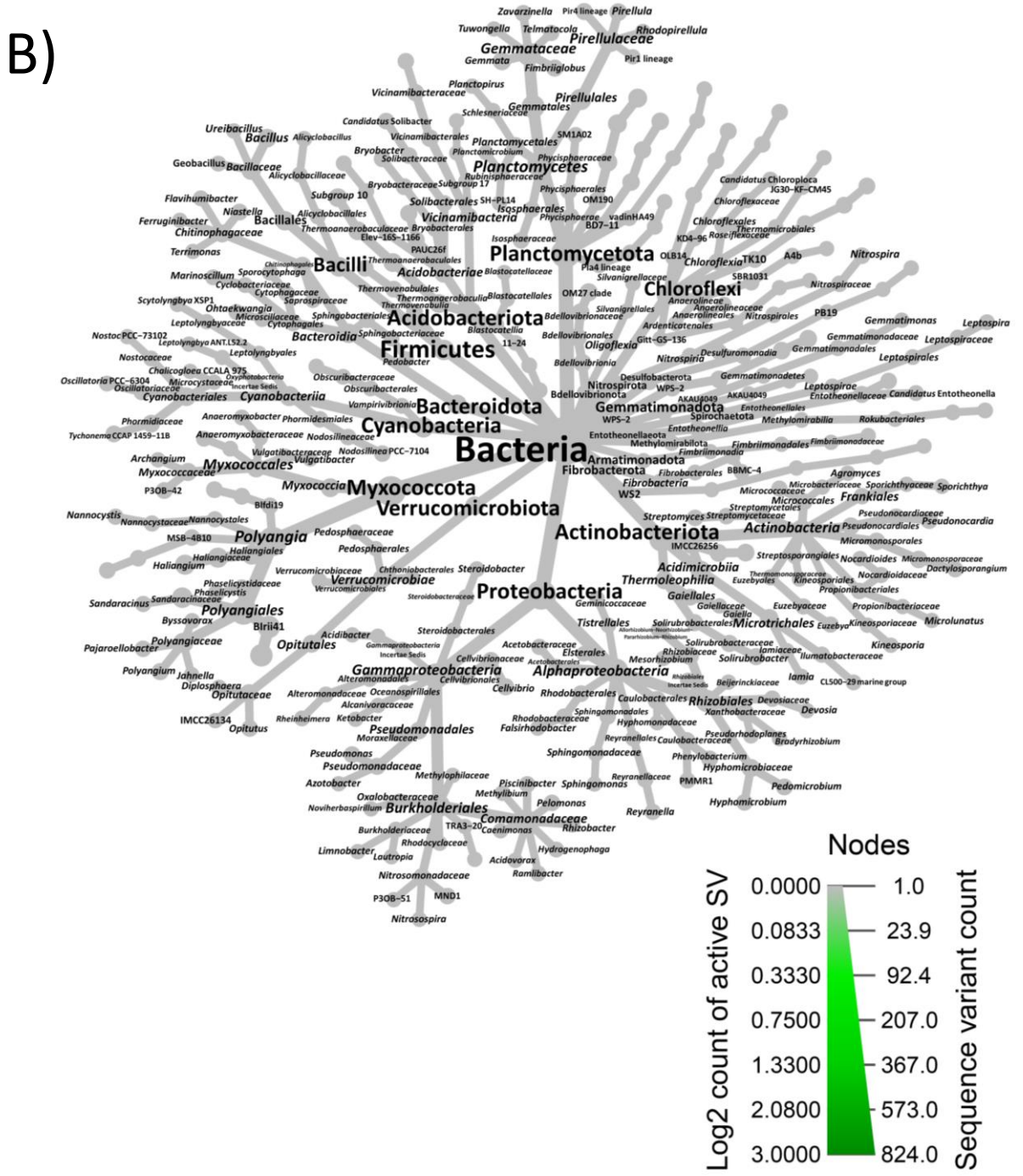


Figure 38 – Taxonomic affiliation of the active sequence variants in bulk soil, the rhizosphere, and inside roots of *H. perforatum, H. olympicum,* and *H. balearicum* in the controlled greenhouse experiment, up to genus level (A). Highlighted with green color are the active bacterial taxa for each group and color intensity reflects the number of active taxa that were classified as that taxon (the values are given in Figure B) while numbers below each tree depict the number of group-active sequence variants for each group. Figure B depicts the key tree with the taxonomic information. Node diameter reflects number of sequence variant classified as that taxon whereas edge width reflects number of reads.

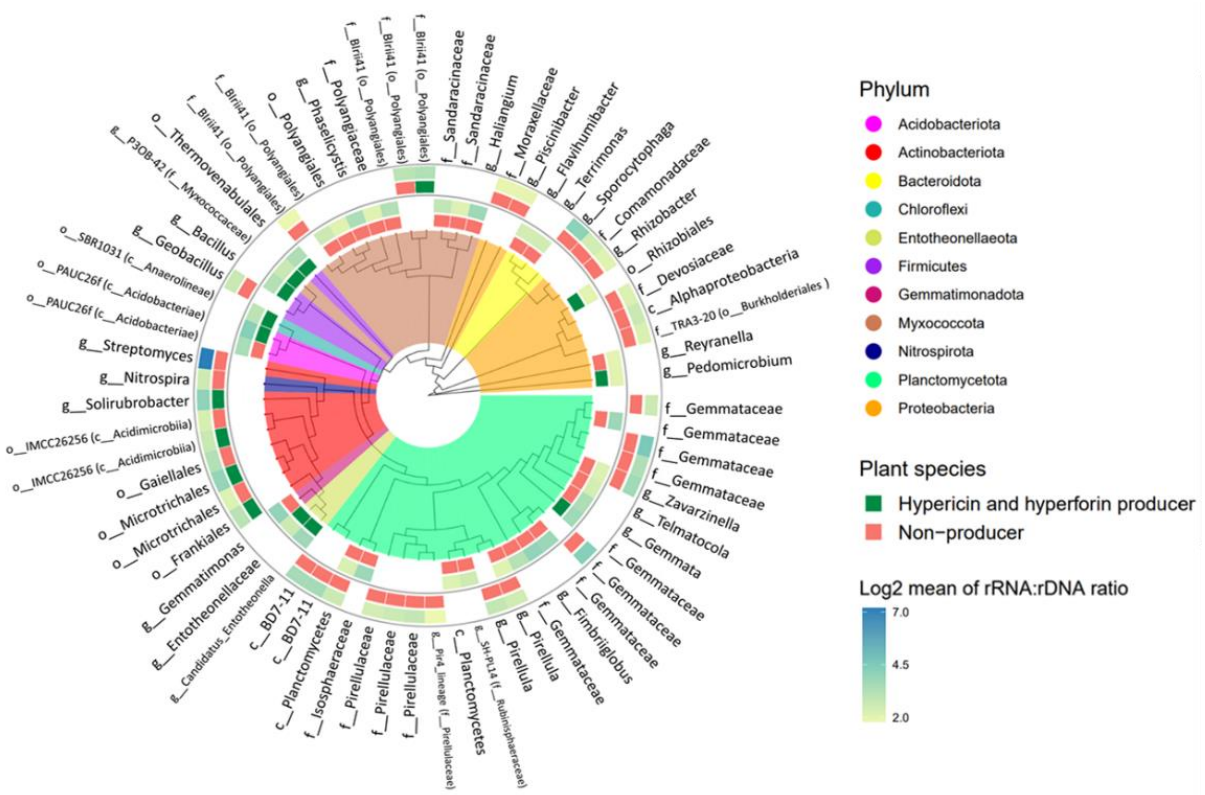


Figure 39 – The specific-active sequence variants in the rhizosphere of hypericin and hyperforin producers (*H. perforatum* and *H. olympicum*) and non-producer (*H. balearicum*) in the controlled greenhouse experiment. Inner two rings represent active taxa on acidic substrate while the outer two rings depict active taxa on alkaline substrate. Ring 1 and 3 (from inside out) represent the plant species (either producer or non-producer) that these taxa are active-specific to and ring 2 and 4 depict the log2 mean of rRNA:RDNA ratio.

The net relatedness index (NRI) and nearest taxon index (NTI) were calculated to evaluate clustering of a particular group-active taxa at deep and terminal branches, respectively. Positive values of NRI and NTI indicate that active taxa are more closely related than expected under the null model while negative values indicate that the taxa are less related. Over the entire tree (not shown), the specific active sequence variants in the rhizosphere of both producer and non-producer, regardless of soil substrate, were phylogenetically underdispersed/clustered at deep and terminal branches (positive NTI and NRI values; Table 4), implying that active taxa of each producer and non-producer are more related than expected under null model. Thus, the result may suggest that there is some sort of plant specificity (between producer and non producer) in selecting their active taxa.

The rhizosphere active taxa that were specific for a particular plant species were also investigated on both acidic and alkaline substrates (Figure 40), as these taxa may play important roles related to the specific ecological functions of the corresponding host plant. The plant species-specific active taxa in

the rhizosphere of the three plant species were distributed across 16 different bacterial phyla, mainly belonged to Planctomycetota, Proteobacteria, Myxococcota, and Actinobacteriota (Figure 40). The NTI and NRI values over the entire tree (not shown) suggested that the active taxa of each plant species are rather closely related, independently of soil substrate, and thus imply plant-specificity on the selection of the active taxa (Table 5).

Table 4 – Net relatedness index (NRI) and nearest taxon index (NTI) values of specific active sequence variants in the rhizosphere of hypericin and hyperforin-producers (*H. perforatum* and *H. olympicum*) and a non-producer (*H. balearicum*). Positive NRI and NTI values indicate that taxa are more closely related than expected under null model while negative values indicate otherwise.

Plant species	NRI		NTI	
	Acidic	Alkaline	Acidic	Alkaline
Hypericin and hyperforin producer	1.7	2.0	1.5	2.0
Non-producer	1.1	1.5	3.5	4.0

Interestingly, members of Planctomycetota were found to be specific-active in the rhizosphere of *H. perforatum* and *H. balearicum* on both acidic and alkaline substrates*,* but not in the rhizosphere of *H. olympicum.* These taxa may contribute to ecological functions that are carried out by both *H. perforatum* and *H. balearicum.* On acidic substrate, members of Actinobacteriota were identified as *H. perforatum* and *H. olympicum* – specific active taxa in contrast to *H. balearicum*. On alkaline substrate, members of Proteobacteria were found to be active mostly in the rhizosphere of *H. balearicum,* with only one sequence variant active in the rhizosphere of *H. olympicum,* suggesting the importance of the phylum to *H. balearicum* plants growing on alkaline condition.

Table 5 – Net relatedness index (NRI) and nearest taxon index (NTI) values of plant species specific-active sequence variants in the rhizosphere of *H. perforatum, H. olympicum,* and *H. balearicum.* Positive NRI and NTI values indicate that taxa are more closely related than expected under null model while negative values indicate otherwise.

Plant species	NRI		NTI	
	Acidic	Alkaline	Acidic	Alkaline
H. perforatum	2.2	1.4	2.4	2.1
H. olympicum	2.9	2.0	2.0	2.2
H. balearicum	1.1	1.5	3.5	4.0

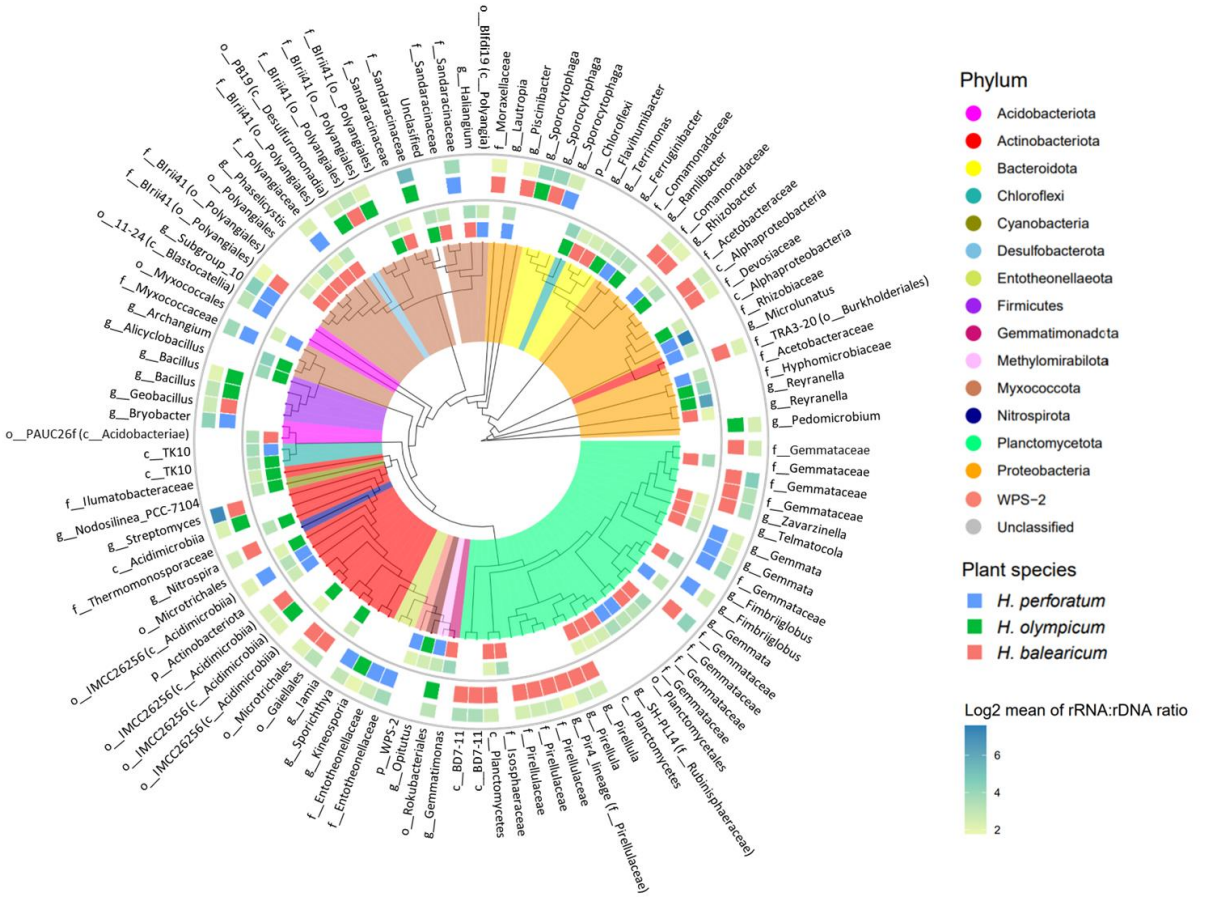


Figure 40 – Plant-species specific active sequence variants in the rhizosphere of *H. perforatum, H. olympicum,* and *H. balearicum* in the controlled greenhouse experiment. Inner two rings represent active taxa on acidic substrate while the outer two rings depict active taxa on alkaline substrate. Ring 1 and 3 (from inside out) represent the plant species that these taxa are active-specific to and ring 2 and 4 depict the log2 mean of rRNA:RDNA ratio.

One particular active sequence variant that belonged to the class *Acidimicrobiia* was consistently identified in *H. perforatum* rhizosphere, regardless of soil substrate. The same pattern was also observed in the rhizosphere of *H. balaearicum* for sequence variants affiliated with *Zavarzinella*, *Pirellula,* and Planctomycetes. Furthermore, one active sequence variant of *H. perforatum* belonging to the family *Rhizobiaceae* and one sequence variant of *H. balearicum* belonging to the genus *Streptomyces,* might be interesting to explore further due to their remarkably high mean values of rRNA:rDNA ratio (191 and 144.8, respectively) in the rhizosphere of their corresponding host plant. These active taxa are likely important and perform functions that are plant species-specific, including elicitation of hypericin and hyperforin in *H. perforatum.* However, the result needs to be validated in the future.

Investigation of specific active sequence variants inside the roots of producer and non-producer was also conducted (excluding the criterium related to the DNA detection across bulk soil samples, as the

active sequence variants inside the roots might be seed-borne and thus not necessarily originated from bulk soil) to find the potential key players in the production of hypericin and hyperforin in the producing-species (Figure 41). The specific active taxa inside the roots of hypericin and hyperforin-producing species belonged to 5 different phyla, namely Proteobacteria, Planctomycetota, Actinobacteriota, Bdellovibrionota, and Myxococcota. The NTI and NRI values imply that active taxa of both producer and non-producer, were phylogenetically underdispersed/clustered at deep and terminal branches over the entire tree (not shown) on both acidic and alkaline substrates (Table 6). Thus, active taxa inside the roots of each group are rather closely related, which may imply plant specificity in choosing their active bacteria.

Table 6 – Net relatedness index (NRI) and nearest taxon index (NTI) values of specific active sequence variants in the roots of hypericin and hyperforin-producers (*H. perforatum* and *H. olympicum*) and a non-producer (*H. balearicum*). Positive NRI and NTI values indicate that taxa are more closely related than expected under null model while negative values indicate otherwise.

Plant species	NRI		NTI	
	Acidic	Alkaline	Acidic	Alkaline
Hypericin and hyperforin producer	3.0	3.4	2.6	3.9
Non-producer	3.2	1.9	6.3	3.8

On acidic substrate, members of *Phenylobacterium*, *Haliangium,* OM27 clade of Bdellovibrionaceae, *Rhizobiaceae,* and *Commamonadaceae* were among the specific-active taxa in the rhizosphere of hypericin and hyperforin-producing species. On alkaline substrate, the specific-active taxa included members of *Rhizobacter, Pseudorhodoplanes, Pedomicrobium, Streptomyces, Sandaracinaceae, Commamonadaceae, Pirellulaceae,* and Blrii41 of *Polyangiales*. Interestingly, one sequence variant belonging to the family *Rhizobiaceae* was consistently active on the rhizosphere of the producers regardless of soil substrates and with high rRNA:rDNA ratio values. Again, the ability of these taxa to trigger hypericin and hyperforin needs to be validated further.

Members of Verrucomicrobiota, Bdellovibrionota, and Spirochaetota were identified as one of the active taxa inside the roots, which were not detected previously for the analysis of the rhizosphere (Figure 39). This may imply that these taxa are more adapted to living condition inside the roots and thus can occupy specific niches inside the habitat. Interestingly, soil substrate seemed to still have an effect in governing the plant-species-specific active taxa inside the roots, albeit the bacterial communities not being in direct contact with it. It is likely that soil substrate affecting plant growth that subsequently lead to changes in the active bacteria inside the roots.

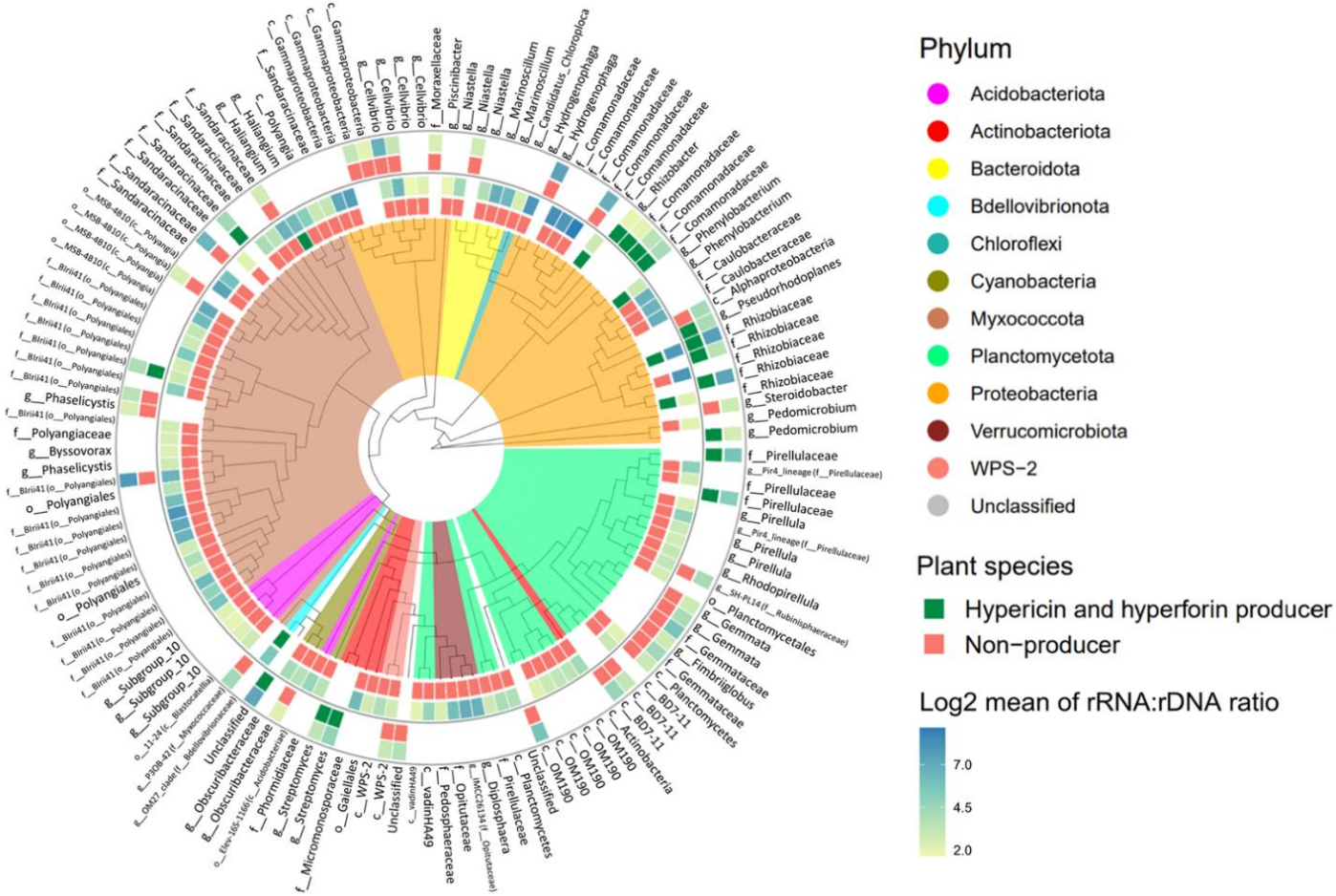


Figure 41 – The specific-active sequence variants inside the roots of hypericin and hyperforin producers (*H. perforatum* and *H. olympicum*) and non-producer (*H. balearicum*) in the controlled greenhouse experiment. Inner two rings represent active taxa on acidic substrate while the outer two rings depict active taxa on alkaline substrate. Ring 1 and 3 (from inside out) represent the plant species (either producer or non producer) that these taxa are active-specific to and ring 2 and 4 depict the log2 mean of rRNA:RDNA ratio.

The active taxa inside the roots that were specific for a particular plant species were also investigated, since they are also potentially important for the corresponding host plant (Figure 42). Similar to what had been observed in the rhizosphere, members of Planctomycetota were mostly observed as the plant species-specific active for *H. balearicum* and *H. perforatum*, with exception of 2 sequence variants that were specifically active on the alkaline substrate of *H. olympicum.* On acidic substrate, 91 sequence variants were identified as *H. balearicum*-specific active taxa while the number of active taxa of *H. olympicum* and *H. perforatum* within this substrate was marginally lower (22 and 23, respectively). On the other hand, 53 and 40 sequence variants were identified as *H. perforatum* and *H. olympicum*-specific active taxa within alkaline substrate, respectively, while the number of *H. balearicum*-specific active taxa within this substrate was only 30. These results imply that the selection of active bacteria by plant is affected by external factors (in this case by distinct soil pH).

Members of the families *Commamonadaceae, Caulobacteriaceae, Rhizobiaceae*, and *Xanthobacteriaceae* had remarkably high mean values of rRNA:rDNA ratio (the lowest value for the last three families is 41.3 while the values ranged from 3.7 to 569 for *Commamonadaceae*), regardless of plant species and soil substrate. These taxa may be crucial for host plant health and productivity in general. In addition, members of *Opitutaceae* and Blrii41 of *Polyangiales* had high rRNA:rDNA ratio in the root-acidic samples of *H. balearicum*. The NRI and NTI values suggest phylogenetic clustering of each group-active taxa in the roots on both acidic and alkaline substrates, at deep and terminal branches (Table 7). Thus, plant-specificity may be evident for the selection of active taxa inside the roots.

Table 7 – Net relatedness index (NRI) and nearest taxon index (NTI) values of plant species specific-active sequence variants inside the roots of *H. perforatum, H. olympicum,* and *H. balearicum*. Positive NRI and NTI values indicate that taxa are more closely related than expected under null model while negative values indicate otherwise.

Plant species	NRI		NTI	
	Acidic	Alkaline	Acidic	Alkaline
H. perforatum	2.1	3.9	4.0	4.8
H. olympicum	3.0	5.2	2.6	5.0
H. balearicum	3.1	1.8	6.3	4.1

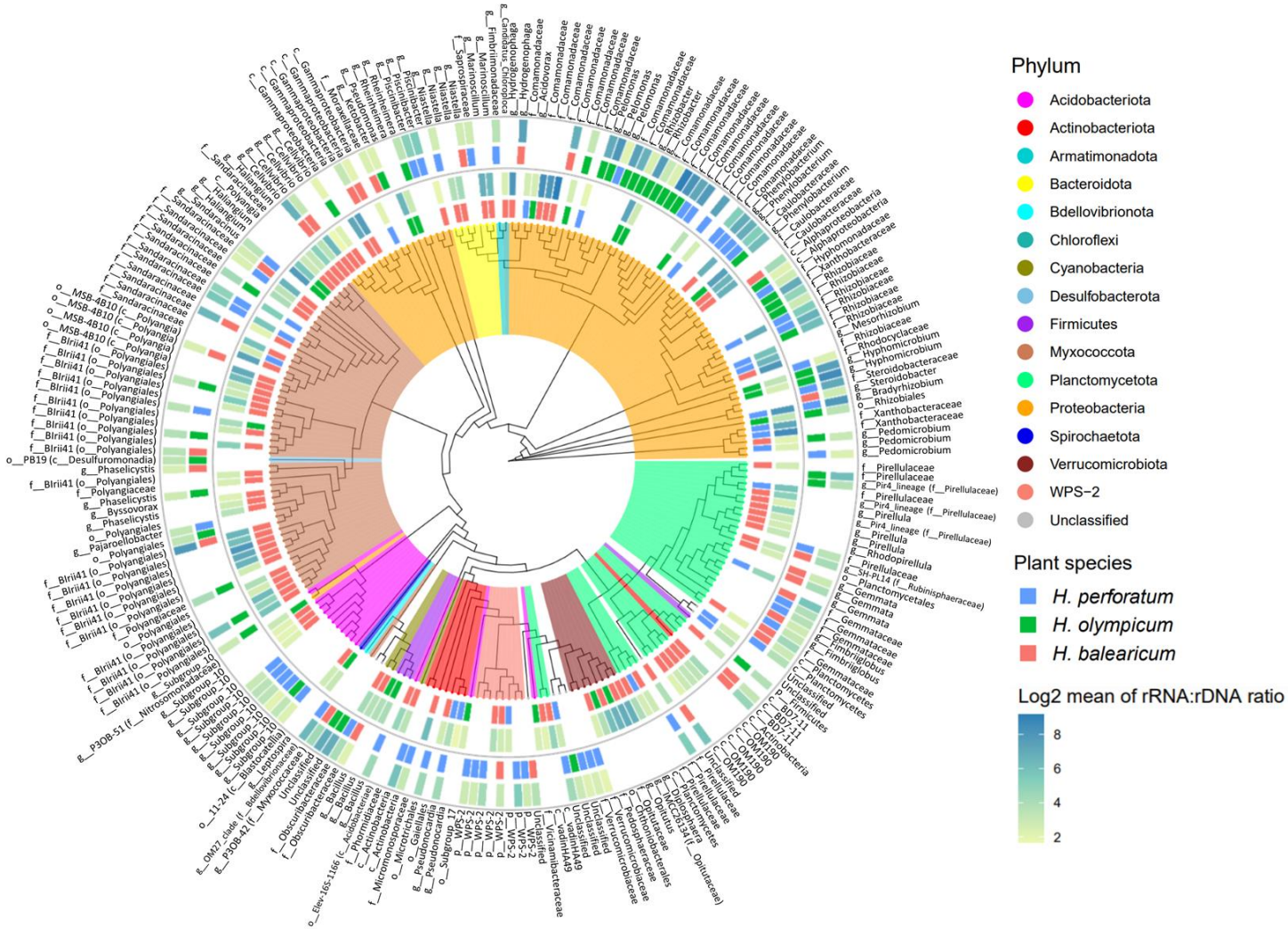


Figure 42 – Plant-species specific active sequence variants inside the roots of *H. perforatum, H. olympicum,* and *H. balearicum* in the controlled greenhouse experiment. Inner two rings represent active taxa on acidic substrate while the outer two rings depict active taxa on alkaline substrate. Ring 1 and 3 (from inside out) represent the plant species that these taxa are active-specific to and ring 2 and 4 depict the log2 mean of rRNA:RDNA ratio.

4.2 – Discussion

As a first step in understanding the connection between bacterial communities associated to *Hypericum* plants and the production of hypericin and hyperforin with anti-anxiety and anti-depression properties, the total (DNA-based) and active (RNA-based) bacterial communities in the rhizosphere and inside roots of several *Hypericum* species were investigated. The impact of habitat type (bulk soil, rhizosphere, and inside roots), soil substrate (acidic and alkaline), and plant species (*H. perforatum, H. olympicum,* and *H. balearicum*) were determined in the controlled greenhouse experiment with the full-factorial crossed design.

Before the controlled greenhouse experiment took place, the investigation of bacterial community of *Hypericum* plants was conducted in prior with plants collected from the Leibniz IPB Halle in 2016 (the first scoping study) and Botanic Garden and Botanical Museum Berlin-Dahlem in 2018 (the second scoping study). The greenhouse samples of Halle, that are associated to the rhizosphere of *H. perforatum*, contain lower richness when compared to other groups. Further investigation confirms the dominance of four sequence variants belonging to the family *Nostococaceae*. The decreased diversity observed in the group is likely due to the heat-sterilization process that was applied to the soils in prior, leading to the elimination of many bacterial taxa residing on the soil. Although members of terrestrial *Nostoc*, such as *N. flagelliforme* and *N. commune*, have a great-heat resistance especially in the dry condition, many soil taxa including the endospore-forming *Clostridium* and *Bacillus* are able to withstand limited heat exposure associated to sterilization process (Sand-Jensen & Jespersen 2012; Sand-Jensen 2014; Nunes *et al.* 2018). Therefore, other selective factors such as high levels of moisture and phosphorus might be responsible for the dominance of *Nostoc* on the soil (Chen *et al.* 2012; Gheda & Ahmed 2015). So far, no report has specifically linked *Nostococaceae* to *H. perforatum.*

Since plants were taken from different locations and that the edaphic properties of the soil are not available, comparison of bacterial communities between different plant species is not possible, as the bulk soil bacterial communities were already different to begin with. Therefore, the comparison is only performed between bulk soil and the rhizosphere of the same plant, as the soil edaphic properties are likely similar and thus differences might be solely based on the plant effect. Samples from the first scoping study (collected from IPB Halle) are completely excluded from the analysis since bulk soil is only represented by low number of biological replicates or completely missing across different plant species. Nonetheless, some samples are still used as the inoculum for high throughput cultivation approach that will be discussed in the following chapter. Moreover, the comparison between the total (DNA-based) and active (RNA-based) bacterial communities is also

conducted, to identify potentially active taxa that may perform specific functions related to the soil or plant ecosystem functioning.

Of RNA-enriched (compared to DNA) and rhizosphere-enriched (compared to bulk soil) taxa, many are known with plant growth promoting rhizobacteria (PGPR) characteristics. These include the known roles of *Streptomyces* and *Pseudomonas* in phosphate solubilization and production of siderophore, indole-3-acetic acid, and antimicrobial compounds (Kaur *et al.* 2019; Qessaoui, R; Suárez-Moreno 2019), *Paenibacillus* in providing protection against plant fungal pathogens (Nielsen & Sørensen 1997), and *Solirubrobacter* in promoting plant growth of apple trees (Franke-Whittle *et al.* 2015).

Furthermore, in accordance with the results of previous studies that emphasize the low percentage of active taxa on the soil (Blagodatskaya 2013; Bowsher *et al.* 2019), approximately 1% of bacterial taxa are identified as active based on a specific threshold on rRNA:rDNA ratio applied in this study, independently of habitat types (bulk soil or rhizosphere), and plant species *(H. perforatum, H. androsaemum,* or *H. balearicum*). Aside from being enriched on RNA and in the rhizosphere, members of *Streptomyces,* and *Pseudomonas* are also determined as active taxa (based on the threshold) in the rhizosphere, independently of plant species. Of other rhizosphere active taxa, *Bacillus* has been reported to promote plant growth by enhancing plant nutrient uptake and inducing plant defense mechanisms against plant pathogen, leading to the utilization as biocontrol agent (Kumar *et al.* 2011). In addition, *Nocardioides* has been reported to degrade mycotoxin deoxynivalenol (DON) that is harmful for crops such as wheat and barley (Ikunaga *et al.* 2011), whereas *Mesorhizobium* promotes nodules formation on chickpea plant (*Cicer arietinum* L.) (Verma *et al.* 2013), and *Phenylobacterium* and *Inquilinus* perform nitrogen fixation (Yang *et al.* 2017). However, the role of other active taxa such as members of *Fimbriiglobus* and *Nakamurella* has not been investigated, which is probably related to the limited number of isolates available in the culture collections (Kulichevskaya *et al.* 2017; Nouioui *et al.* 2017). Therefore, the roles of these active taxa to the host plant must be further investigated, by improving the cultivation success or applying activity-based study such as metatranscriptomic and RNA-based stable isotope probing (RNA-SIP).

This investigation provides an overview of bacterial communities of wild and cultivated *Hypericum* plants growing in Leibniz IPB Halle and BGBM Berlin. The analysis is hindered by issues such as the differences in the bulk soil bacterial communities associated to distinct sampling locations, missing information on soil edaphic properties, and insufficient numbers of biological replicates. In order to

deal with the issues, a controlled-greenhouse experiment with full-factorial crossed design was executed in 2019 in BGBM Berlin, which will be discussed in the following sub-chapter.

4.2.1 – Habitat, soil substrate, and plant species contribute to differences on *Hypericum* bacterial communities

Many factors have been reported to govern soil bacterial communities, including pH (Kaiser *et al.* 2016), organic carbon (Shen *et al.* 2015), oxygen availability, soil texture and structure (Fierer 2017), moisture (Freedman and Zak 2015), nitrogen and phosphorus content (Shen *et al.* 2015), and vegetation type (Zhang *et al.* 2014). In addition, the study of bacterial colonization on minerals has revealed that *de novo* mineral surfaces drive the communities when the minerals were exposed to grassland soils, exceeding the effect of carbon sources type and environmental factors (Vieira *et al.* 2020a). For plant-associated microbiomes, habitat (rhizosphere, roots, phyllosphere, *etc.*) and root exudates are among the strongest drivers (Haichar *et al.* 2008, Berg *et al.* 2014). Investigation of the active bacterial communities of 6 plant species in temperate grasslands of central Europe has revealed that the recruitment from bulk soil governs the rhizosphere communities, and that the effect of soil variables exceeds those of plant species and root exudates composition (Vieira *et al.* 2020b).

In this controlled greenhouse experiment with full-factorial crossed design, habitat is the strongest driver of bacterial community structure, followed by soil substrate and plant species. Out of the three habitats (bulk soil, rhizosphere and roots), the root harbour significantly lower alpha diversity measures, which is likely due to the necessary prerequisites to colonize root such as motility, biofilm production and adaptation to enter the root, along with the ability of the bacteria to survive inside the living tissues and maintain stronger interactions than outside of the root (Minz *et al.* 2013; Frank *et al.* 2017; Afzal *et al.* 2019).

Lower numbers of reads are observed on the root samples when compared to bulk soil and rhizosphere, owing to the removal of many chloroplast and mitochondrial reads in prior. Thus, a more advanced technique should be applied to optimize the recovery of these reads in the coming study of root endophytic communities. Almost no significant differences can be observed between bulk soil and rhizosphere samples, which is in accordance with many preceding studies (Rathore *et al.* 2017; Luo *et al.* 2017; Lee *et al.* 2019). A reduced species richness has been commonly reported in the rhizosphere when compared to bulk soil, mainly due to selective factors such as root exudation and a reduced niche dimension in the former habitat (Peiffer *et al.* 2013; Shi *et al.* 2015; Fernández-Gómez; Garcia-Lemos 2020). On the contrary, an increase on the richnesss of the active

(RNA-based) bacterial communities in the rhizosphere when compared to bulk soil has been reported in two out of three regions in European temperate grasslands (Vieira *et al.* 2020b). The authors argue that the presence of relics DNA and dormant taxa might contributes to higher richness on bulk soil in preceeding studies that are mostly based on DNA and thus not observed in RNA-based study. For the active bacterial communities (RNA-based) of *Hypericum*, the same pattern is observed only on alkaline substrate of *H. perforatum.* A recent study has revealed that the bias in biodiversity estimates increases with relic DNA pool size (Lennon *et al.* 2018). However, this only happen when the species abundance distribution (SAD) of relic DNA differ from the intact DNA. It is argued that even in the ecosystem that promotes the presence of relic DNA (e.g. rich in biofilm), relic DNA may not alter biodiversity estimates if the SAD of relic DNA is similar to the intact DNA. The similarity in species richness between the DNA-based and RNA-based communities in our study may suggest that the SAD of relic and intact DNA are not distinct from one another across all samples, and thus bias due to relic DNA is almost negligible.

Investigation of habitat-specific communities revealed that soil substrate, with distinct pH values, strongly drive bulk soil communities. Moreover, this study also confirms the effect of plant species in shaping the rhizosphere and root communities (contributes to approximately 20% of variance), which has been reported previously. The other reported plant species effect to be either higher, such as depicted by one third of variance explained by the effect of the plant in differentiating bulk soil and rhizosphere communities of *Ranunculus glacialis* distributed along the altitudinal gradient in the Alphine; Praeg *et al.* 2019, or lower with 4.4% and 0.86% of variance explained in the permafrost regions of Northeastern China and in the main tailings following ecological restoration with *Imperata cylindrica* and *Chrysopogon zizanioides* plants, respectively (Li *et al.* 2016; Ren *et al.* 2018). In addition, soil characteristics (30% of variance) dominate over plant species (2%) in the study of active bacterial communities in temperate grasslands that has been mentioned previously (Vieira *et al.* 2020b). It is argued that the interaction between rhizosphere of different plant species masked the effect of plant species in such a complex natural habitat like grasslands. Therefore, for a controlled greenhouse experiment where the same standard soils are employed and each individual plant is cultivated in a separate pot (as employed in this study), the effect of plant species is likely more pronounced.

Interestingly, soil substrate effect is still pronounced in driving the root endophytic communities despite the fact that root habitat is not in direct contact with soil substrate itself. Further observation has revealed that the majority of root taxa are also found on bulk soil and the rhizosphere. Since the root communities seem to derive from bulk soil communities and that root

endophytes can be transmitted either vertically from seedlings or horizontally from the surrounding soil (Frank *et al.* 2017), the effect of soil substrate on bulk soil communities might leads to changes on the root communities. The assembly of the root bacterial communities depends on different factors. For soil taxa, the colonialization success depends on plant-associated factors, such as rhizodeposition and the influence of plant to soil pH and structures, and oxygen content in the root area (Bulgarelli *et al.* 2013; Philippot *et al.* 2013; Frank *et al.* 2017), and bacteria-associated factors, including motility, chemotaxis, quorum-sensing, and biofilm production that is beneficial for root attachment (De Weger *et al.* 1987; Wei & Zhang 2006; Edwards *et al.* 2015; Scharf *et al.* 2016; Frank *et al.* 2017). Moreover, bacteria can enter the roots either by producing endoglucanase and endopolygalacturonase that degrade the plant cell walls (Compant *et al.* 2005; Frank *et al.* 2017) or by direct entry via cracks or emerging lateral roots (Liu *et al.* 2017).

The plant species effect is evident for bulk soil communities. However, the impact is almost negligible and separation based on plant species is hardly noticeable on the ordination plot. Although the roots of *H. balearicum* appeared to be small in size, this is not the case for *H. olympicum* and *H. perforatum*. Root exudates, in the form of dissolved organic carbon, is able to diffuse up to 28 mm distance from the root surface, which is the maximum sampling distance employed in the reported-experiment and thus might travel further (Hafner *et al.* 2014). Since plants are cultivated in small pots and some roots are quite high in density, root exudates might diffuse relatively far from roots and hence the observed plant species effect on bulk soil communities.

Taxonomic shift of the communities between the total (DNA-based) and active (RNA-based) bacterial communities, and between distinct soil substrate, is quite pronounced. However, less taxa are differentially abundant between the two soil substrates and between the two community types (total or active) for root habitat when compared to bulk soil and the rhizosphere. Of all detected genera on bulk soil, rhizosphere, and roots, approximately 72%, 44%, and 34% are differentially abundant between DNA and RNA data, respectively. Moreover, around 80%, 61% and 43% are enriched either on acidic or alkaline substrate for bulk soil, rhizosphere, and roots, respectively. These smaller numbers of differentially abundant-taxa between distinct community type and between distinct soil substrate on the roots compared to other two habitats might be explained by significant numbers of taxa that are missing in the root, as also supported by lower alpha diversity measures.

Differences between total (RNA-based) and active (DNA-based) bacterial communities are prominent for all three habitats, which imply different level of activity across the observed taxa (Blagodatskaya & Kuzyakov 2013; Fierer 2017). *Piscinibacter, Phenylobacterium,* and *Polyangium* are

among those that are enriched in the active (RNA-based) communities when compared to the total ones (DNA-based), independently of habitat. Although information about activity of these taxa on soil is lacking, *Piscinibacter* is abundant in the rhizosphere of rice paddies (Zecchin *et al.* 2017) and able to reduce nitrate (Chen *et al.* 2018), while *Phenylobacterium* is a known diazotroph (Yang *et al.* 2017) and *Polyangium* produces antimicrobial compound (Masschelein *et al.* 2017).

Nitrospira, Reyranella, Euzebya, and *Ornithinimicrobium* are among the enriched genera in the active (RNA-based) bacterial communities when compared to the total ones (DNA-based), specifically on the rhizosphere. Of these, *Nitrospira* has recently gained significant interest and termed as comammox, due to the active role in executing complete oxidation of ammonia to nitrate in soils (Li *et al.* 2019; Wang *et al.* 2019), while members of *Reyranella* is able to reduce nitrate (Pagnier *et al.* 2011). Moreover, although their roles on the rhizosphere has yet been identified, some members of *Ornithinimicrobium*, such as *O. pekingense* and *O. kibberense*, have a wide range of substrates that they can utilize (Fang *et al.* 2020). Thus, their enrichment in the active (RNA-based) bacterial communities may reflect the ability to utilize various substrates available in the soil at that particular time.

The enriched genera in the active (RNA-based) bacterial communities inside the roots include *Opitutus, Ilumatobacter*, *Rhizocola,* PMMR1 of *Caulobacteraceae*, and subgroup 17 of Acidobacteriota. Of these, *Opitutus* has been reported to grow on plant-derived polysaccharides (van Passel *et al.* 2011). The subgroup 17 abundance might be corelated with S, K, and C element in soil but there is no information available for the role of the taxon inside the roots (Navarrete *et al.* 2015; de Chaves 2019). Moreover, the role of *Ilumatobacter* and *Rhizocola* related to their host plant is still not investigated albeit detected in the roots of salt marsh plants and *Helleborus orientalis*, respectively (Matsumoto *et al.* 2014; Chen *et al.* 2019).

The enriched-taxa on alkaline substrate when compared to acidic substrate, regardless of community type (total or active), plant species and habitat, include *Piscinibacter, Pseudomonas,* and *Niastella*. Members of these genera, such as *Piscinibacter caeni, Piscinibacter defuvii* and *Niastella gongjuensis* have been reported to grow on alkaline conditions with pH values of approximately 9 (Cho *et al.* 2016; Yan *et al.* 2016; Chen *et al.* 2018), while *Pseudomonas alcaliphila* and *Pseudomonas toyotomiensis* have been reported as alkaliphiles (Yumoto *et al.* 2001; Hirota *et al.* 2010). On the other hand, the enriched taxa on acidic substrate include *Sphingomonas, Mesorhizobium, Pseudolabrys,* and *Nocardioides,* but these taxa are not necessarily considered as acidophiles (Kämpfer *et al.* 2006; Kämpfer *et al.* 2016; Dludlu *et al.* 2018; Xu *et al.* 2020). Their enrichment on acidic substrate might reflects the contribution of some acidophilic members that are yet-

uncultured. This also implies that the ability to grow on distinct soil substrates, with different pH, varies across members of the same genera.

Based on variance partitioning analysis employed for each distinct habitat, the plant species impact dominates over soil substrate on the rhizosphere and root bacterial communities. The plant species affect bacterial community by modulating the litter quality and root exudates profiles (Whitham *et al.* 2003; Purahong *et al.* 2016). Moreover, the influence of the plant on the bacterial communities is assumed to be related to their ability in modifying soil pH, structures, and oxygen availability (Marschner *et al.* 1986; Dennis *et al.* 2010; Frank *et al.* 2017). The higher impact of plant species compared to soil substrate on the rhizosphere and root communities in this study is in accordance with the study on the root areas of broadleaf tree species, where the impact of tree species identity exceeds soil pH on the soil enzyme activity (Purahong *et al.* 2016). However, it has been commonly reported that soil physicochemical effect dominates over plant species on the rhizosphere microbial communities (Navrátilová *et al.* 2018, Ren *et al.* 2018; Habiyaremye *et al.* 2020), although recent report suggests that the effect is mainly explained by soil structure, water content, and soil type (Vieira *et al.* 2020). The more pronounced impact of plant species on the rhizosphere and roots communities on this study may be related to the high density of roots of *H. perforatum* and *H. olympicum* and the relatively small pots used to grow the plants. In this case, it might be easier for the plants to strongly affect the rhizosphere bacterial communities, as also supported by the fact that the plant effect is detected even on bulk soil bacterial communities.

Taxa enrichment is evident between distinct plant species based on the result of pairwise comparison. *H. perforatum*-enriched taxa should be investigated further as the host plant is the most notable species for the production of hypericin and hyperforin with anti-anxiety and anti-depression effects (Kusari *et al.* 2009). On the other hand, *H. balearicum* is considered a non-producer (Stojanović 2013). Several studies confirm *H. olympicum* as a hypericin and hyperforin producer (Kitanov 2001; Smelcerovic 2006; Stojanović 2013), while other described the plant as a non-producer (Crockett *et al.* 2005). Since the concentration of hypericin and hyperforin in *H. olympicum* is lower compared to *H. perforatum* (Smelcerovic 2006; Kusari *et al.* 2015), it might be below the detection limit on some cases and thus not detected. For that reason, we considered *H. olympicum* as a producer of both compounds.

Aeromicrobium, Microlunatus, Bryobacter, and *Xenophilus* are consistently enriched in the rhizosphere of *H. perforatum* only in RNA data, regardless of the paired plant species. RNA data is emphasized since it represents the active bacterial communities. These genera might modulate the concentration of *Hypericum* secondary metabolites (including hypericin and hyperforin), since it has

been reported that beneficial soil bacteria are able to trigger plant secondary metabolite through induction of systemic resistance (ISR) (Harun-Or-Rashid & Chung 2017). This resistance mechanism is part of the plant defense against pathogens and herbivores, where microbe-associated molecular patterns (MAMPs) of beneficial microbes, such as flagellin, secondary metabolites, lipopolysaccharides (LPS), and exopolysaccharides (EPS) are recognized by the plant, leading to the activation of jasmonic acid (JA), ethylene (ET), or salicylic acid pathways (SA) (Zamioudis & Pieterse 2012; Harun-Or-Rashid & Chung 2017). It has been reported that root colonization by *P. simiae* WCS417r trigger JA/ET pathways that lead to ISR against chewing insects (Pangesti *et al.* 2016) while lipopeptides produced by *B. amyloliquefaciens* S499 induce ISR in tomato plants, by triggering the expression of the defense-related genes lipoxygenase D and F (Cawoy *et al.* 2014).

Although *H. perforatum*-enriched taxa own PGPR characteristics, including *Aeromicrobium* with the production of indole-3-acetic acid (IAA), siderophores, and antibiotics (Yadav *et al.* 2015; Pan & Chu 2016), *Microlunatus* and *Bryobacter* with growth promotion of durum wheat shoot (*Triticum turgidum* L. var. *durum* (Desf.) Husn.) and wild *Marchantia*, respectively (Yang *et al.* 2013; Alcaraz *et al.* 2018), and *Xenophilus* with nitrogen fixation and improvement of plant stress tolerance (Singh *et al.* 2015; Deng *et al.* 2020), their role in modulating the plant secondary metabolites are unknown. Therefore, the ability of these taxa to trigger hypericin and hyperforin in *H. perforatum* should be investigated further. So far, only *Pseudomonas fluorescens* and *Stenotrophomonas maltophilia* have been known to increase hypericin and pseudohypericin content when inoculated into the seedlings of *H. perforatum* (Mañero *et al.* 2012).

Investigation of the consistently-enriched taxa on the hypericin and hyperforin-producing species (*H. perforatum* and *H. olympicum*) when compared to a non-producer (*H. balearicum*) might provide initial knowledge related to the potential taxa that elicit hypericin and hyperforin. However, these potential taxa cannot be fully determined in this study since the production of hypericin and hyperforin was confounded by plant species. In order to validate the result and to further identify bacterial taxa that trigger hypericin and hyperforin, a deeper investigation of a specific plant producer grown in different soil substrates (for example acidic, neutral, and alkaline soils) could be executed, where the content of hypericin and hyperforin is measured directly. The bacterial community can then be investigated in order to find taxa with increasing activity as the content of hypericin and hyperforin in the plant increases.

Some of the consistently-enriched taxa on the hypericin and hyperforin-producing species are considered diazotroph (*Panacagrimonas* and *Azohydromonas*; Xie & Yokota 2005; Li *et al.* 2018), sulphate reducing bacteria (*Desulfonatronum* and *Desulfosporosinus*; González-Toril 2011;

Ryzhmanova *et al.* 2013; Ayangbenro *et al.* 2018), and biofilm-producing bacteria (*Adhaeribacter;* McBride *et al.* 2014), while others play a role in P cycling (*Panacagrimonas*; Wu *et al.* 2020), plant materials degradation (Calleja-Cervantes *et al.* 2015; Pii *et al.* 2016), production of phytohormone, and siderophore (*Elioraea*; Lee & Whang 2020) and improvement of plant salt stress tolerance (*Oceanobacillus*; Bensidhoum & Nabti 2019). Although plant growth-promoting characteristics of some taxa have been reported, the link between these taxa and the plant secondary metabolites remains unclear. Some taxa including *Panacagrimonas* are considered plant pathogen and thus might trigger hypericin and hyperforin through PAMP-triggered immunity (PTI), where pathogen-associated molecular patterns are recognized by the plant and leads to the activation of defense signaling mechanisms (Pieterse *et al.* 2009).

4.2.2 – A small fraction of the taxa are active across distinct bulk soil or rhizosphere groups

According to the specific threshold based on rRNA:rDNA ratio applied in this study, approximately 1% of taxa are active in bulk soil and the rhizosphere, with roots contain a slightly higher percentage (up to 2.5%). This small fraction of active taxa on soil has been observed in preceeding studies employing different methods such as direct microscopy with cell staining, RNA-based FISH combined with staining of total biomass, and microbial growth-based approaches (Blagodatskaya & Kuzyakov 2013). Therefore, the threshold is reliable to identify active taxa, since the ouput is in line with previous findings. The rest of the taxa are in a dormant state, which act as a seed bank and can be active given favorable environmental conditions (Khomutova *et al.* 2003; Yarwood *et al.* 2013; Wang *et al.* 2014; Bickel & Or 2020). In addition, across all groups and especially observed inside the roots, the numbers of active taxa are higher on alkaline substrate when compared to acidic. This result is quite interesting since plants have lower biomass yield on alkaline substrates, implying a higher level of stress on the substrate. Thus, a higher percentage of active taxa on alkaline substrate might be related to the role of bacteria in mitigating plant stress due to alkaline condition. Lower plant biomass due to alkaline condition has been reported in a study investigating the response of rhizosphere bacteria of *Cucumis sativus* to salinity, pH, and boron (Ibekwe *et al.* 2010). In addition, *Agrobacterium* and *Bacillus* have been reported to improve plant survival under alkaline condition by promoting vegetative growth and nutrient uptake (Karaca & Sabir 2018), enhancing seed germination and photosynthetic pigments, and reducing proline accumulation to increase plant stress tolerance (Dar *et al.* 2016; Dixit *et al.* 2020).

Of bulk soil active taxa, soil-substrate specificity is pronounced as shown with the enrichment of OM27 clade of Bdellovibrionota, *Telmatocola,* TK10 of Chloroflexi, *Sphingomonas,* and *Dactylosporangium* on acidic substrate and the enrichment of *Pseudomonas, Methylibium,*

Fimbriiglobus, and *Euzebya* on alkaline substrate. On the other hand, soil substrate-specificity is almost negligible inside the roots. Lower impact of soil pH on the root bacterial community compared to the rhizosphere has been reported previously (Li *et al.* 2016). The results imply that the activity of relevant taxa is limited by pH, and hence the observed lower impact of soil substrate on the root active taxa, since here pH is most likely constant. The pH limitation for specific taxa has also been reported, such as with the study of *Pseudomonas protegens*CHAo that revealed the loss of culturability of the strain on acidic soil while growth is retained on nonacidic soils (Mascher *et al.* 2014) and the acidophilic nature of members of Planctomycetota, including *Telmatocola* (Dedysh & Kulichevskaya 2013).

Interestingly, some taxa are active in all rhizosphere groups, independently of soil substrate and plant species, including members of *Streptomyces* and *Phenylobacterium*. Within the roots, these include members of *Phenylobacterium, Ohtaekwangia,* and *Haliangium.* The specific roles of the mentioned genera to the host plant have been previously reported, such as the production of haliangicin by *Haliangium* that inhibits the growth of plant fungal pathogens (Fudou *et al.* 2001; Wang *et al.* 2012) or the ability to modulate *Eucalyptus* secondary metabolites of *Streptomyces* (Salla *et al.* 2014). In addition, *Phenylobacterium* is considered diazotroph (Yang *et al.* 2017) and able to degrade pesticide (You *et al.* 2016). Although the role of *Ohtaekwangia* in the rhizosphere is still unknown, enrichment of this bacterium in the rhizosphere is positively correlated with higher apple yield (Wang *et al.* 2016).

In addition to the effect against anxiety and depression, hypericin and hyperforin also harbour neuroprotective activity against Alzheimer disease (Griffith *et al.* 2010; Hofrichter *et al.* 2013; Rizzo *et al.* 2019). Hypericin and hyperforin can be elicited by jasmonic acid and salicylic acid, respectively, confirming their involvement in plant defense mechanisms against plant pathogen or herbivorous insect (Sirvent & Gibson 2002; Sirvent *et al.* 2003). Correspondingly, jasmonic acid/ethylene (JA/ET) and salicylic acid (SA) pathways can be activated by beneficial microbes and pathogens in the rhizosphere of the plant species, respectively (Harun-Or-Rashid & Chung 2017). In addition, while lacking hypericin, root cultures of *H. perforatum* contain hyperforin (Gaid *et al.* 2016; Gaid *et al.* 2019), implying the possibility of the compound to be released as root exudates that subsequently select the beneficial bacteria. Therefore, the active sequence variants in the rhizosphere of hypericin and hyperforin producers should be investigated further since they might be specifically recruited and play crucial roles in the elicitation of the plant secondary metabolites.

The active sequence variants in the rhizosphere of hypericin and hyperforin-producing species are phylogenetically diverse. On acidic substrate, specific-active taxa in the rhizosphere of the producer

(*H. perforatum* and *H. olympicum*) include members of *Bacillus* and *Pedomicrobium*. Although the ability to trigger hypericin and hyperforin is yet-unknown for *Bacillus*, *Bacillus cereus* AR156 is able to trigger both JA/ET and salicylic acid-responsive marker genes (PDF1.2 and PR1, PR2, and PR5, respectively) in the leaves of *Arabidopsis thaliana* when the plant is exposed to *P. syringae* (Niu *et al.* 2011). Although less information is available for *Pedomicrobium,* the genus is abundant in some soil type (Alcaraz *et al.* 2018) and forms biofilm (Sly *et al.* 1988). Since the rhizosphere also harbour rhizoplane communities due to the root-washing step with detergent prior to rhizosphere collection, *Pedomicrobium* might belong to the rhizoplane community where biofim formation is quite advantageous for the root attachment.

Since *H. perforatum* is widely accepted as the most valuable source of hypericin and hyperforin, also when compared to *H. olympicum* (Smelcerovic 2006; Kusari *et al.* 2015), H. *perforatum*-specific active taxa are further investigated. Although members of *H. perforatum*-specific active taxa have yet been reported to trigger plant secondary metabolites, some of them including *Rhizobacter, Sporocytophaga, Ramlibacter,* and *Sporichthya,* are considered PGPR (Mohanram & Kumar 2019; Zhang *et al.* 2019).

Root (endophytic) bacterial communities have the potential to trigger the production of plant secondary metabolites via induced systematic resistance (ISR) (Pieterse *et al.* 2014; Harun-Or-Rashid & Chung 2017), although it is not proven so far for *Hypericum*. Nonetheless, the specific-active taxa inside the roots of hypericin and hyperforin producer are investigated further. Of those active taxa, endophytic *Pseudonocardia* has been reported to elicit artemisinin in *Artemisia annua* (Li *et al.* 2014). In addition, *Streptomyces, Mesorhizobium,* and both *Bacillus* and *Bradyrhizobium* are involved in induced systemic resistance that trigger secondary metabolites of oak (*Quercus robur* L.), chickpea (*Cicer arietinum* L.), and *Arabidopsis thaliana*, respectively (Cartieaux *et al.* 2008; Niu *et al.* 2011; Kurth *et al.* 2014; Kumari & Khanna 2020).

One active sequence variant belonging to the family *Rhizobiaceae* is observed with high rRNA:DNA ratio in the acidic-rhizosphere samples of *H. perforatum*. Since the taxon is only identified at family level, molecular methods such metatrascriptomics should be employed to identify specific functions that the bacterium carries in the rhizosphere of *H. perforatum*. Correspondingly, many other active taxa are only classified at higher taxonomic level, including members of *Comamonadaceae* and *Steroidobacteraceae*. This would also assist in cultivation efforts to obtain such key players triggering the plant unique secondary metabolites as pure cultures.

The present study confirms the effect of plant species and soil substrate on the bacterial communities in the rhizosphere and roots of *H. perforatum*, *H. olympicum,* and *H. balearicum*. In addition, the structure and composition of the total (DNA-based) and active (RNA-based) bacterial communities are not the same in bulk soil, rhizosphere and roots, implying different level of activities across bacterial taxa on each habitat. We further identified active taxa across all groups based on the rRNA:rDNA ratio values, and postulate the potential ecological role of active taxa associated to the hypericin and hyperforin-producing species. However, the activity of these taxa can also reflect the ability to utilize different carbon sources that are available in the soil or roots at that particular time and not necessarily related to the production of secondary metabolites. Therefore, a comprehensive study needs to be performed in the future to reveal the specific ecological functions that these active taxa carry in the rhizosphere or inside the roots of *Hypericum* plants.

4.3 – References

Afzal, I., Shinwari, Z. K., Sikandar, S. & Shahzad, S. 2019. Plant beneficial endophytic bacteria: Mechanisms, diversity, host range and genetic determinants. *Microbiological Research*, **221**, 36-49, doi: 10.1016/j.micres.2019.02.001.

Alcaraz, L.D., Peimbert, M., Barajas, H.R., Dorantes-Acosta, A. E., Bowman, J. L. & Arteaga-Vázquez, M. A. 2018. *Marchantia* liverworts as a *proxy* to plants' basal microbiomes. *Scientific Reports*, **8**, doi: 10.1038/s41598-018-31168-0.

Ayangbenro, A. S., Olanrewaju, O. S. & Babalola, O. O. 2018. Sulfate-Reducing Bacteria as an Effective Tool for Sustainable Acid Mine Bioremediation. *Frontiers in Microbiology,* **9**, doi: 10.3389/fmicb.2018.01986.

Bensidhoum, L. & Nabti, E. 2019. Plant Growth-Promoting Bacteria for Improving Crops Under Saline Conditions. *In:* Giri B., Varma A. (eds) *Microorganisms in Saline Environments: Strategies and Functions.* Springer, Cham, Soil Biology, vol 56, doi: 10.1007/978-3-030-18975-4_14.

Berg, G., Grube, M., Schloter, M. & Smalla, K. 2014. Unraveling the plant microbiome: looking back and future perspectives. *Frontiers in Microbiology*, **5**, doi: 10.3389/fmicb.2014.00148.

Bickel, S. & Or, D. 2020. Soil bacterial diversity mediated by microscale aqueous-phase processes across biomes. *Nature Communications*, **11**, doi: 10.1038/s41467-019-13966-w.

Blagodatskaya, E. & Kuzyakov, Y. 2013. Active microorganisms in soil: Critical review of estimation criteria and approaches. *Soil biology and biochemistry*, **67**, 192-211, doi: 10.1016/j.soilbio.2013.08.024.

Bowsher, A. W., Kearns, P. J. & Shade, A. 2019. The 16S rRNA/rRNA Gene Ratios and Cell Activity Staining Reveal Consistent Patterns of Microbial Activity in Plant-Associated Soil. *mSystems*, **4**, doi: 10.1128/mSystems.00003-19.

Bulgarelli, D., Schlaeppi, K., Spaepen, S., van Themaat, E. V. L. & Schulze-Lefert, P. 2013. Structure and Functions of the Bacterial Microbiota of Plants. *Annual Review of Plant Biology*, **64**, 807-838, doi: 10.1146/annurev-arplant-050312-120106.

Cartieaux, F., Contesto, C., Gallou, A., Desbrosses, G., Kopka, J., Taconnat, L., Renou, J. -P. & Touraine, B. 2008. Simultaneous Interaction of *Arabidopsis thaliana* with *Bradyrhizobium* Sp. Strain ORS278 and *Pseudomonas syringae* pv. *tomato* DC3000 Leads to Complex Transcriptome Changes. *Molecular Plant-Microbe Interactions*, **21**, 244-259, doi: 10.1094/MPMI-21-2-0244.

Cawoy, H., Mariutto, M., Henry, G., Fisher, C., Vasilyeva, N., Thonart, P., Dommes, J. & Ongena, M. 2014. Plant defense stimulation by natural isolates of *Bacillus* depends on efficient surfactin production. *Molecular Plant-Microbe Interactions*, **27**, 87-100, doi: 10.1094/MPMI-09-13-0262-R.

Calleja-Cervantes, M. E., Menéndez, S., Fernández-González, A. J., Irigoyen, I., Cibriáin-Sabalza, J. F., Toro, N., Aparicio-Tejo, P. M. & Fernández-López, M. 2015. Changes in soil nutrient content and bacterial community after 12 years of organic amendment application to a vineyard. *European Journal of Soil Science*, **66**, doi: 10.1111/ejss.12261.

Chen, D. -Z., Yu, N. -N., Chu, Q. -Y, Chen, J., Ye, J. -X., Cheng, Z. -W., Zhang, S. -H. & Chen, J. -M. 2018. *Piscinibacter caeni* sp. nov., isolated from activated sludge. **International Journal of Systematic and Evolutionary Microbiology**, **68**, 2627-2632, doi: 10.1099/ijsem.0.002891.

Chen, P., Zhang, C., Ju, X., Xiong, Y., Xing, K. & Qin, S. 2019. Community Composition and Metabolic Potential of Endophytic Actinobacteria From Coastal Salt Marsh Plants in Jiangsu, China. *Frontiers in Microbiology*, **10**, doi: 10.3389/fmicb.2019.01063.

Chen, Z., Chen, S., Lu, G. & Chen, X. 2012. Phosphorus limitation for the colony formation, growth and photosynthesis of an edible cyanobacterium, *Nostoc sphaeroides*. *Biotechnology Letters*, **34**, 137–143, doi: 10.1007/s10529-011-0747-9.

Cho, S. H., Lee, H. J. & Jeon, C. O. 2016. *Piscinibacter defluvii* sp. nov., isolated from a sewage treatment plant, and emended description of the genus *Piscinibacter*. *International Journal of Systematic and Evolutionary Microbiology*, **66**, 4839-4843, doi: 10.1099/ijsem.0.001438.

Compant, S., Kaplan, H., Sessitsch, A., Nowak, J., Ait Barka, E. & Clément, C. 2008. Endophytic colonization of *Vitis vinifera* L. by *Burkholderia phytofirmans* strain PsJN: From the rhizosphere to inflorescence tissues. *FEMS Microbiology Ecology*, **63**, 84-93, doi: 10.1111/j.1574-6941.2007.00410.x.

Crockett, S. L., Schaneberg, B. & Khan, I. A. 2005. Phytochemical profiling of new and old world *Hypericum* (St. John's Wort) species. *Phytochemical Analysis*, **16**, 479-485, doi: 10.1002/pca.875.

Dar, M. I., Naikoo, M. I., Rehman, F., Naushin, F. & Khan, F. A. 2016. Proline Accumulation in Plants: Roles in Stress Tolerance and Plant Development. *In:* Iqbal N., Nazar R., A. Khan N. (eds) *Osmolytes and Plants Acclimation to Changing Environment: Emerging Omics Technologies*. Springer, New Delhi., doi: 10.1007/978-81-322-2616-1_9.

de Chaves, M. G., Silva, G. G. Z., Rossetto, R., Edwards, R. A., Tsai, S. M. & Navarrete, A. A. 2019. Acidobacteria Subgroups and Their Metabolic Potential for Carbon Degradation in Sugarcane Soil Amended With Vinasse and Nitrogen Fertilizers. *Frontiers in Microbiology*, **10**, doi: 10.3389/fmicb.2019.01680.

De Weger, L. A., van der Vlugt, C. I., Wijfjes, A. H., Bakker, P.A., Schippers, B. & Lugtenberg, B. 1987. Flagella of a plant-growth-stimulating *Pseudomonas fluorescens* strain are required for colonization of potato roots. *Journal of Bacteriology*, **169**, 2769-2773, doi: 10.1128/jb.169.6.2769-2773.1987.

Dedysh S.N. & Kulichevskaya I.S. 2013. Acidophilic Planctomycetes: Expanding the Horizons of New Planctomycete Diversity. *In:* Fuerst J. (eds) *Planctomycetes: Cell Structure, Origins and Biology.* Humana Press, Totowa, NJ., doi: 10.1007/978-1-62703-502-6_5.

Dennis, P. G., Miller, A. J. & Hirsch, P. R. 2010. Are root exudates more important than other sources of rhizodeposits in structuring rhizosphere bacterial communities? *FEMS Microbiology Ecology*, **72**, 313-327, doi: 10.1111/j.1574-6941.2010.00860.x.

Deng, Z. -S., Kong, Z. -Y., Zhang, B. -C. & Zhao, L. -F. 2020. Insights into non-symbiotic plant growth promotion bacteria associated with nodules of *Sphaerophysa salsula* growing in northwestern China. *Archives of Microbiology*, **202**, 399-409, doi: 10.1007/s00203-019-01752-7.

Dixit, V. K., Misra, S., Mishra, S. K., Tewari, S. K., Joshi, N. & Chauhan, P. S. 2020. Characterization of plant growth-promoting alkalotolerant *Alcaligenes* and *Bacillus* strains for mitigating the alkaline stress in *Zea mays*. *Antonie van Leeuwenhoek*, **113**, 889-905, doi: 10.1007/s10482-020-01399-1.

Dludlu, M. N., Chimphango, S. B. M., Stirton, C. H. & Muasya, A. M. 2018. Differential Preference of *Burkholderia* and *Mesorhizobium* to pH and Soil Types in the Core Cape Subregion, South Africa. *Genes*, **9**, doi: 10.3390/genes9010002.

Edwards, J., Johnson, C., Santos-Medellín, C., Lurie, E., Podishetty, N. K., Bhatnagar, S., Eisen, J. A. & Sundaresan, V. 2015. Structure, variation, and assembly of the root-associated microbiomes of rice. *Proceedings of the National Academy of Sciences*, **112**, E911–E920, doi: 10.1073/pnas.1414592112.

Fang, X. -M. Du, H. -J., Bai, J. -L., He, W. -N., Li, J., Wang, H., Su, J., Liu, H. -Y. Zhang, Y. -Q. & Yu, L. -Y. 2020. *Ornithinimicrobium cerasi* sp. nov., isolated from the fruit of *Cerasus pseudocerasus* and emended description of the genus *Ornithinimicrobium*. *International Journal of Systematic and Evolutionary Microbiology*, **70**, doi: 10.1099/ijsem.0.003957.

Fernández-Gómez, B., Maldonado, J., Mandakovic, D., Gaete, A., Gutiérrez, R., Maass, A., Cambiazo, V. & González, M. 2019. Bacterial communities associated to Chilean altiplanic native plants from the Andean grasslands soils. *Scientific Reports*, **9**, doi: 10.1038/s41598-018-37776-0.

Fierer, N. 2017. Embracing the unknown: disentangling the complexities of the soil microbiome. *Nature Reviews Microbiology*, **15**, 579–590, doi: 10.1038/nrmicro.2017.87.

Frank, A. C., Saldierna-Guzmán, J. P., Shay, J. E. 2017. Transmission of Bacterial Endophytes. *Microorganisms*, **5**, doi: 10.3390/microorganisms5040070.

Franke-Whittle, I. H., Manici, L. M., Insam, H. & Stres, B. 2015. Rhizosphere bacteria and fungi associated with plant growth in soils of three replanted apple orchards. *Plant and Soil*, **395**, 317-333, doi: 10.1007/s11104-015-2562-x.

Freedman, Z. & Zak, D. R. 2015. Soil bacterial communities are shaped by temporal and environmental filtering: evidence from a long-term chronosequence. *Environmental Microbiology*, **17**, doi: 10.1111/1462-2920.12762.

Fudou, R., Iizuka, T. & Yamanaka, S. 2001. Haliangicin, a novel antifungal metabolite produced by a marine myxobacterium. 1. Fermentation and biological characteristics. *The Journal of* Antibiotics, **54**, 149-152, doi:10.7164/antibiotics.54.149.

Gaid, M., Grosch, J. -H., Möller, S., Beerhues L. & Krull, R. 2019. Toward enhanced hyperforin production in St. John's wort root cultures. *Engineering in Life Sciences*, **19**, 916-930, doi: 10.1002/elsc.201900043.

Gaid, M., Haas, P., Beuerle, T., Scholl, S. & Beerhues, L. 2016. Hyperforin production in *Hypericum perforatum* root cultures. *Journal of Biotechnology*, **222**, 47-55, doi: 10.1016/j.jbiotec.2016.02.016.

Garcia-Lemos, A. M., Gobbi, A., Nicolaisen, M. H., Hansen, L. H., Roitsch, T., Veierskov, B. & Nybroe, O. 2020. Under the Christmas Tree: Belowground Bacterial Associations with *Abies nordmanniana* Across Production Systems and Plant Development. *Frontiers in Microbiology*, **11**, doi: 10.3389/fmicb.2020.00198.

Gheda, S. F. & Ahmed, D. A. 2015. Improved soil characteristics and wheat germination as influenced by inoculation of *Nostoc kihlmani* and *Anabaena cylindrica*. *Rendiconti Lincei*, **26**, 121–131, doi: 10.1007/s12210-014-0351-8.

González-Toril, E., Aguilera, Á., Souza-Egipsy, V., Pamo, E. L., España, J. S. & Amils, R. 2011. Geomicrobiology of La Zarza-Perrunal Acid Mine Effluent (Iberian Pyritic Belt, Spain). *Applied and Environmental Microbiology,* **77**, 2685-2694, doi: 10.1128/AEM.02459-10.

Griffith, T. N., Varela-Nallar, L., Dinamarca, M. C. & Inestrosa, N. C. 2010. Neurobiological effects of Hyperforin and its potential in Alzheimer's disease therapy. *Current Medicinal Chemistry*, **17**, 391-406, doi: 10.2174/092986710790226156.

Habiyaremye, J. d. D., Goldmann, K., Reitz, T., Herrmann, S. & Buscot, F. 2020. Tree Root Zone Microbiome: Exploring the Magnitude of Environmental Conditions and Host Tree Impact. *Frontiers in Microbiology*, **11**, doi: 10.3389/fmicb.2020.00749.

Haichar, F. Z., Marol, C., Berge, O., Rangel-Castro, J. I., Prosser, J. I., Balesdent, J., Heulin, T. & Achouak, W. 2008. Plant host habitat and root exudates shape soil bacterial community structure. *The ISME Journal*, **2**, 1221–1230, doi: 10.1038/ismej.2008.80.

Hafner, S., Wiesenberg, G. L. B., Stolnikova, E., Merz, K. & Kuzyakov, Y. 2014. Spatial distribution and turnover of root-derived carbon in alfalfa rhizosphere depending on top- and subsoil properties and mycorrhization. *Plant and Soil*, **380**, 101-115, doi: 10.1007/s11104-014-2059-z.

Harun-Or-Rashid, Md. & Chung, Y. R. Induction of Systematic Resistance against Insect Herbivores in Plants by Beneficial Soil Microbes. 2017. *Frontiers in Plant Science*, **8**, doi: 10.3389/fpls.2017.01816.

Hirota, K., Yamahira, K., Nakajima, K., Nodasaka, Y., Okuyama, H. & Yumoto, I. 2011. *Pseudomonas toyotomiensis* sp. nov., a psychrotolerant facultative alkaliphile that utilizes hydrocarbons. *International Journal of Systematic and Evolutionary Microbiology*, **61**, 1842-1848, doi: 10.1099/ijs.0.024612-0.

Hofrichter, J., Krohn, M., Schumacher, T., Lange, C., Feistel, B., Walbroel, B., Heinze, H. J., Crockett, S., Sharbel, T. F. & Pahnke, J. 2013. Reduced Alzheimer's disease pathology by St. John's Wort treatment is independent of hyperforin and facilitated by ABCC1 and microglia activation in mice. *Current Alzheimer Research*, **10**, 1057-1069, doi: 10.2174/15672050113106660171.

Ibekwe, A. M., Poss, J. A., Grattan, S. R., Grieve, C. M. & Suarez, D. 2010. Bacterial diversity in cucumber (*Cucumis sativus*) rhizosphere in responseto salinity, soil pH, and boron. *Soil Biology and Biochemistry*, **42**, 567-575, doi: 10.1016/j.soilbio.2009.11.033.

Ikunaga, Y., Sato, I., Grond, S., Numaziri, N., Yoshida, S., Yamaya, H., Hiradate, S., Hasegawa, M., Toshima, H., Koitabashi, M., Ito, M., Karlovsky, P. & Tsushima, S. 2011. *Nocardioides* sp. strain WSN05-2, isolated from a wheat field, degrades deoxynivalenol, producing the novel intermediate 3-*epi*-deoxynivalenol. *Applied Microbiology and Biotechnology*, **89**, 419-427, doi: 10.1007/s00253-010-2857-z.

Kaiser, K., Wemheuer, B., Korolkow, V., Wemheuer, F., Nacke, H., Schöning, I., Schrumpf, M. & Daniel, R. 2016. Driving forces of soil bacterial community structure, diversity, and function in temperate grasslands and forests. Scientific Reports, **6**, doi: 10.1038/srep33696.

Kämpfer, P., Glaeser, S. P., McInroy, J. A. & Busse, H. -J. 2016. *Nocardioides zeicaulis* sp. nov., an endophyte actinobacterium of maize. *International Journal of Systematic and Evolutionary Microbiology*, **66**, 1869-1874, doi: 10.1099/ijsem.0.000959.

Kämpfer, P., Young, C. -C., Arun, A. B., Shen, F. -T., Jäckel, U., Rosselló-Mora, R., Lai, W. -A. & Rekha, P. D. 2006. *Pseudolabrys taiwanensis* gen. nov., sp. nov., an alphaproteobacterium isolated

from soil. *International Journal of Systematic and Evolutionary Microbiology*, **56**, 2469-2472, doi: 10.1099/ijs.0.64124-0.

Karaca, U. & Sabir, A. 2018. Sustainable Mitigation of Alkaline Stress in Grapevine Rootstocks (*Vitis* spp.) by Plant Growth-Promoting Rhizobacteria. *Erwerbs-Obstbau*, **60**, 211–220, doi: 10.1007/s10341-017-0361-7.

Kaur, M., Jangra, M., Singh, H., Tambat, R., Singh, N., Jachak, S. M., Mishra, S., Sharma, C., Nandanwar, H. & Pinnaka, A. K. 2019. *Pseudomonas koreensis* Recovered from Raw Yak Milk Synthesizes a β-Carboline Derivative with Antimicrobial Properties. *Frontiers in Microbiology*, **10**, doi: 10.3389/fmicb.2019.01728.

Khomutova, T. E., Demkina, T. S. & Demkin, V. A. 2004. Estimation of the Total and Active Microbial Biomasses in Buried Subkurgan Paleosoils of Different Age. *Microbiology*, **73**, 196–201, doi: 10.1023/B:MICI.0000023989.04745.7b.

Kitanov, G. M. 2001. Hypericin and pseudohypericin in some *Hypericum* species. *Biochemical Systematics and Ecology*, **29**, 171–178, doi: 10.1016/S0305-1978(00)00032-6.

Kulichevskaya, I. S., Ivanova, A. A., Baulina, O. I., Rijpstra, W. I. C., Sinninghe Damsté, J. S. & Dedysh, S. N. 2017. *Fimbriiglobus ruber* gen. nov., sp. nov., a *Gemmata*-like planctomycete from *Sphagnum* peat bog and the proposal of *Gemmataceae* fam. nov. *International Journal of Systematic and Evolutionary Microbiology*, **67**, 218-224, doi: 10.1099/ijsem.0.001598.

Kumar, A., Prakash, A. & Johri, B. N. 2011. *Bacillus* as PGPR in Crop Ecosystem. *In:* Maheshwari D. (eds) *Bacteria in Agrobiology: Crop Ecosystems.* Springer Berlin, Heidelberg., doi: 10.1007/978-3-642-18357-7_2.

Kumari, S. & Khanna, V. 2020. Induction of Systemic Resistance in Chickpea (*Cicer arietinum* L.) Against *Fusarium oxysporum* f. sp. *ciceris* by Antagonistic Rhizobacteria in Assistance with Native *Mesorhizobium*. *Current Microbiology*, **77**, 85-98, doi: 10.1007/s00284-019-01805-6.

Kurth, F., Mailänder, S., Bönn, M., Feldhahn, L., Herrmann, S., Große, I., Buscot, F., Schrey, S. D. & Tarkka, M. T. 2014. *Streptomyces*-Induced Resistance Against Oak Powdery Mildew Involves Host Plant Responses in Defense, Photosynthesis, and Secondary Metabolism Pathways. *Molecular Plant-Microbe Interactions*, **27**, 891-900, doi: 10.1094/MPMI-10-13-0296-R.

Kusari, S., Sezgin, S., Nigutova, K., Cellarova, E. & Spiteller, M. 2015. Spatial chemo-profiling of hypericin and related phytochemicals in *Hypericum* species using MALDI-HRMS imaging. *Analytical and Bioanalytical Chemistry*, **407**, 4779-4791, doi: 10.1007/s00216-015-8682-6.

Kusari, S. Zühlke, S., Borsch, T. & Spiteller, M. 2009. Positive correlations between hypericin and putative precursors detected in the quantitative secondary metabolite spectrum of Hypericum. *Phytochemistry*. **70**, 1222-1232. doi: 10.1016/j.phytochem.2009.07.022.

Lee, H. -J. & Whang, K. -S. 2020. *Elioraea rosea* sp. nov., a plant promoting bacterium isolated from floodwater of a paddy field. *International Journal of Systematic and Evolutionary Microbiology*, **70**, doi: 10.1099/ijsem.0.004028.x.

Lennon, J. T., Muscarella, M. E., Placella, S. A. & Lehmkuhl, B. K. 2018. How, When, and Where Relic DNA Affects Microbial Diversity. *MBio*, **9**, e00637-18, doi: 10.1128/mBio.00637-18.

Li, C., Hu, H., Chen, Q., Chen, D., & He, J. 2019. Comammox *Nitrospira* play an active role in nitrification of agricultural soils amended with nitrogen fertilizers. *Soil Biology & Biochemistry*, **138**, doi: 10.1016/j.soilbio.2019.107609.

Li, D., Voigt, T. B. & Kent, A. D. 2016. Plant and soil effects on bacterial communities associated with *Miscanthus* x *giganteus* rhizosphere and rhizomes. *Global Change Biology Bioenergy*, **8**, 183-193, doi: 10.1111/gcbb.12252.

Li, J., Zhao, G. -Z., Varma, A., Qin, S., Xiong, Z., Huang, H. -Y., Zhu, W. -Y., Zhou, L. -X., Xu, L. -H., Zhang, S. & Li, W. -J. 2012. An Endophytic *Pseudonocardia* Species Induces the Production of Artemisinin in *Artemisia annua*. *PLoS ONE*, **7**, doi: 10.1371/journal.pone.0051410.

Li, Y., Jia, Z., Sun, Q., Zhan, J., Yang, Y. & Wang, D. 2016. Ecological restoration alters microbial communities in mine tailings profiles. *Scientific Reports*, **6**, doi: 10.1038/srep25193.

Li, Y., Wu, X., Chen, T., Wang, W., Liu, G., Zhang, W., Li, S., Wang, M., Zhao, C., Zhou, H. & Zhang, G. 2018. Plant Phenotypic Traits Eventually Shape Its Microbiota: A Common Garden Test. *Frontiers in Microbiology*, **9**, doi: 10.3389/fmicb.2018.02479.

Liu, H., Carvalhais, L. C., Crawford, M., Singh, E., Dennis, P. G., Pieterse, C. M. J. & Schenk, P. M. 2017. Inner Plant Values: Diversity, Colonization and Benefits from Endophytic Bacteria. *Frontiers in Microbiology*, **8**, doi: 10.3389/fmicb.2017.02552

Lee, S. A., Kim, Y., Kim, J. M., Chu, B., Joa, J. -H., Sang, M. K., Song, J. & Weon H. -Y. 2019. A preliminary examination of bacterial, archaeal, and fungal communities inhabiting different rhizocompartments of tomato plants under real-world environments. *Scientific Reports,* **9**, doi**:** 10.1038/s41598-019-45660-8.

Lu, T., Ke, M., Lavoie, M., Jin, Y., Fan, X., Zhang, Z., Fu, Z., Sun, L., Gillings, M., Peñuelas, J., Qian, H. & Zhu, Y. -G. 2018. Rhizosphere microorganisms can influence the timing of plant flowering. *Microbiome*, **6**, doi: 10.1186/s40168-018-0615-0.

Luo, J., Tao, Q., Wu, K., Li, J., Qian, J., Liang, Y., Yang, X. & Li, T. 2017. Structural and functional variability in root-associated bacterial microbiomes of Cd/Zn hyperaccumulator *Sedum alfredii*. *Applied Microbiology and Biotechnology*, **101,** 7961-7976, doi: 10.1007/s00253-017-8469-0.

Mañero, F. J. G., Algar, E., Gómez, M. S. M., Sierra, M. D. S. & Solano, B. R. 2012. Elicitation of secondary metabolism in Hypericum perforatum by rhizosphere bacteria and derived elicitors in seedlings and shoot cultures. *Pharmaceutical Biology*, **50**, 1201-1209, doi: 10.3109/13880209.2012.664150.

Marschner, H., Römheld, V., Horst, W. J. & Martin, P. 1986. Root-induced changes in the rhizosphere: Importance for the mineral nutrition of plants. *Journal of Plant Nutrition and Soil Science*, **149**, 441-456, doi: 10.1002/jpln.19861490408.

Mascher, F., Hase, C., Bouffaud, M. -L., Défago, G. & Moënne-Loccoz, Y. 2014. Cell culturability of *Pseudomonas protegens*CHA0 depends on soil pH. *FEMS Microbiology Ecology,* **87**, 441–450, doi: 10.1111/1574-6941.12234.

Masschelein, J., Jenner, M. & Challis, G. L. 2017. Antibiotics from Gram-negative bacteria: a comprehensive overview and selected biosynthetic highlights. *Natural Product Reports*, **34**, 712-783, doi: 10.1039/c7np00010c.

Matsumoto, A., Kawaguchi, Y., Nakashima, T., Iwatsuki, M., Ōmura, S. & Takahashi, Y. 2014. *Rhizocola hellebori* gen. nov., sp. nov., an actinomycete of the family *Micromonosporaceae* containing 3,4-dihydroxydiaminopimelic acid in the cell-wall peptidoglycan. *International Journal of Systematic and Evolutionary Microbiology*, **64**, 2706-2711, doi:10.1099/ijs.0.060293-0.

McBride, M. J., Liu, W., Lu, X., Zhu, Y. & Zhang, W. 2014. The family Cytophagaceae. *In: The Prokaryotes*: *Other Major Lineages of Bacteria and The* Archaea. 577-593., doi: 10.1007/978-3-642-38954-2_382.

Minz, D., Ofek, M. & Hadar, Y. 2013. Plant Rhizosphere Microbial Communities. *In*: Rosenberg E., DeLong E.F., Lory S., Stackebrandt E., Thompson F. (eds) *The Prokaryotes*. Springer, Berlin, Heidelberg, doi: 10.1007/978-3-642-30123-0_38.

Mohanram, S., Kumar, P. 2019. Rhizosphere microbiome: revisiting the synergy of plant-microbe interactions. *Annals of Microbiology*, **69**, 307-320, doi: 10.1007/s13213-019-01448-9.

Navarrete, A. A., Diniz, T. R., Braga, L. P., Silva, G. G., Franchini, J. C., Rossetto, R., Edwards, R. A. & Tsai, S. M. 2015. Multi-Analytical Approach Reveals Potential Microbial Indicators in Soil for Sugarcane Model Systems. *PLoS One*, **10**, doi: 10.1371/journal.pone.0129765.

Navrátilová, D., Tláskalová, P., Kohout, P., Dřevojan, P., Fajmon, K., Chytrý, M. & Baldrian, P. 2019. Diversity of fungi and bacteria in species-rich grasslands increases with plant diversity in shoots but not in roots and soil. *FEMS Microbiology Ecology*, **95**, doi: 10.1093/femsec/fiy208.

Nielsen, P. & Sørensen, J. 1997. Multi-target and medium-independent fungal antagonism by hydrolytic enzymes in *Paenibacillus polymyxa* and *Bacillus pumilus* strains from barley rhizosphere. *FEMS Microbiology Ecology*, **22**, 183-192, doi: 10.1111/j.1574-6941.1997.tb00370.x.

Niu, D. -D., Liu, H. -X., Jiang, C. -H., Wang, Y. -P., Wang, Q. -Y, Jin, H. -L. & Guo, J. -H. 2011. The plant growth–promoting rhizobacterium *Bacillus cereus* AR156 induces systemic resistance in *Arabidopsis thaliana* by simultaneously activating salicylate and jasmonate/ethylene-dependent signaling pathways. *Molecular Plant-Microbe Interactions*, **24**, 533-542, doi: 10.1094/MPMI-09-10-0213.

Nouioui, I., Göker, M., Carro, L., Montero-Calasanz, M. D. C., Rohde, M., Woyke, T., Kyrpides, N. C. & Hans-Peter Klenk, H. -P. 2017. High quality draft genome of *Nakamurella lactea* type strain, a rock actinobacterium, and emended description of *Nakamurella lactea*. *Standards in Genomic Sciences*, **12**, doi: 10.1186/s40793-016-0216-0.

Nunes, I., Jurburg, S., Jacquiod, S., Brejnrod, A., Falcão Salles, J., Priemé, A. & Sørensen, S. J. 2018. Soil bacteria show different tolerance ranges to an unprecedented disturbance. *Biology and Fertility of Soils*, **54**, 189-202, doi: 10.1007/s00374-017-1255-4.

Pan, M. & Chu, L. M. 2016. Phytotoxicity of veterinary antibiotics to seed germination and root elongation of crops. *Ecotoxicology & Environmental Safety*, **126**, 228-237, doi: 10.1016/j.ecoenv.2015.12.027.

Pangesti, N., Reichelt, M., van de Mortel, J. E., Kapsomenou, E., Gershenzon, J., van Loon, J. J., Dicke, M. & Pineda, A. 2016. Jasmonic acid and ethylene signaling pathways regulate glucosinolate levels in plants during rhizobacteria-induced systemic resistance against a leaf-chewing herbivore. *Journal of Chemical Ecology*, **42**, 1212-1225.doi: 10.1007/s10886-016-0787-7.

Pagnier, I., Raoult, D., & La Scola, B. 2011. Isolation and characterization of *Reyranella massiliensis* gen. nov., sp. nov. from freshwater samples by using an amoeba co-culture procedure. *International Journal of Systematic and Evolutionary Microbiology*, **61**, 2151–2154, doi: 10.1099/ijs.0.025775-0.

Peiffer, J. A., Spor, A., Koren, O., Jin, Z., Tringe, S. G., Dangl, J. L., Buckler, E. S. & Ley, R. E. 2013. Diversity and heritability of the maize rhizosphere microbiome under field conditions. *Proceedings of the National Academy of Sciences*, **110**; 6548-53; doi: 10.1073/pnas.1302837110.

Philippot, L., Raaijmakers, J.M., Lemanceau, P. & van der Putten, W. H. 2013. Going back to the roots: The microbial ecology of the rhizosphere. *Nature Reviews Microbiology*, **11**, 789-799, doi: 10.1038/nrmicro3109.

Pieterse, C. M. J., Zamioudis, C., Berendsen, R. L., Weller, D. M., Van Wees, S. C. M. & Bakker, P. A. H. M. 2014. Induced Systemic Resistance by Beneficial Microbes*. Annual Review of Phytopathology*, 52, 347-375, doi: 10.1146/annurev-phyto-082712-102340.

Pieterse, C., Leon-Reyes, A., Van der Ent, S. & Van Wees, S. C. M. 2009. Networking by small-molecule hormones in plant immunity. *Nature Chemical Biology*, **5**, 308-316, doi: 10.1038/nchembio.164.

Pii, Y., Borruso, L., Brusetti, L. Crecchio, C., Cesco, S. & Mimmo, T. 2016. The interaction between iron nutrition, plant species and soil type shapes the rhizosphere microbiome. *Plant Physiology and Biochemistry*, **99**, 39-48, doi: 10.1016/j.plaphy.2015.12.002.

Praeg, N., Pauli, H. & Illmer, P. 2019. Microbial Diversity in Bulk and Rhizosphere Soil of *Ranunculus glacialis* Along a High-Alpine Altitudinal Gradient. *Frontiers in Microbiology*, **10**, doi: 10.3389/fmicb.2019.01429.

Purahong, W., Durka, W., Fischer, M., Dommert, S., Schöps, R., Buscot, F. & Wubet, T. 2016. Tree species, tree genotypes and tree genotypic diversity levels affect microbe-mediated soil ecosystem functions in a subtropical forest. *Scientific Reports*, **6,** doi: 10.1038/srep36672.

Qessaoui, R., Bouharroud, R., Furze, J.N. Aalaoui, M. El., Akroud, H., Amarraque, A., Van Vaerenbergh, J., Tahzima, R., Mayad, E. H. & Chebli, B. 2019. Applications of New Rhizobacteria *Pseudomonas* Isolates in Agroecology via Fundamental Processes Complementing Plant Growth. *Scientific Reports*, **9**, doi: 10.1038/s41598-019-49216-8.

Rabbee, M. F., Ali, M. S., Choi, J., Hwang, B. S., Jeong, S. C. & Baek, K. -H. 2019. *Bacillus velezensis:* A Valuable Member of Bioactive Molecules within Plant Microbiomes. *Molecules* **2019**, *24*, 1046.

Rathore, R., Dowling, D. N., Forristal, P. D., Spink, J., Cotter, P. D., Bulgarelli, D. & Germaine, K. J. 2017. Crop Establishment Practices Are a Driver of the Plant Microbiota in Winter Oilseed Rape (*Brassica napus*). *Frontiers in Microbiology*, **8**, doi: 10.3389/fmicb.2017.01489.

Ren, B., Hu, Y., Chen, B., Zhang, Y., Thiele, J., Shi, R., Liu, M. & Bu, R. 2018. Soil pH and plant diversity shape soil bacterial community structure in the active layer across the latitudinal gradients in continuous permafrost region of Northeastern China. *Scientific Reports*, **8**, doi**:** 10.1038/s41598-018-24040-8.

Rizzo, P., Altschmied, L., Stark, P., Rutten, T., Gündel, A., Scharfenberg, S., Franke, K., Bäumlein, H., Wessjohan, L., Koch, M., Borisjuk, L. & Sharbel, T. F. 2019. Discovery of key regulators of dark gland development and hypericin biosynthesis in St. John's Wort (*Hypericum perforatum*). *Plant Biotechnol Journal,* **17**, 2299-2312, doi: 10.1111/pbi.13141.

Ryzhmanova, Y., Nepomnyashchaya, Y., Abashina, T., Ariskina, E., Troshina, O., Vainshtein, M. & Shcherbakova, V. 2013. New sulfate-reducing bacteria isolated from Buryatian alkaline brackish lakes: description of *Desulfonatronum buryatense* sp. nov. *Extremophiles*, **17**, 851-859, doi: 10.1007/s00792-013-0567-z.

Salla, T. D., da Silva, T. R., Astarita, L. V. & Santarém, E. R. 2014. *Streptomyces* rhizobacteria modulate the secondary metabolism of *Eucalyptus* plants. *Plant Physiology and Biochemistry*, **85**, 14-20, doi: 10.1016/j.plaphy.2014.10.008.

Sand-Jensen, K. & Jespersen, T. S. 2012. Tolerance of the widespread cyanobacterium *Nostoc commune* to extreme temperature variations (-269 to 105°C), pH and salt stress. *Oecologia*, **169**, 331-339, doi: 10.1007/s00442-011-2200-0.

Sand-Jensen, K. 2014. Ecophysiology of gelatinous *Nostoc* colonies: unprecedented slow growth and survival in resource-poor and harsh environments. *Annals of Botany*, **114**, 17–33, doi: 10.1093/aob/mcu085**.**

Scharf, B. E., Hynes, M. F. & Alexandre, G. M. 2016. Chemotaxis signaling systems in model beneficial plant-bacteria associations. *Plant Molecular Biology*, **90**, 549-559, doi: 10.1007/s11103-016-0432-4.

Shen, C, Ni, Y., Liang, W., Wang, J. & Chu, H. 2015. Distinct soil bacterial communities along a small-scale elevational gradient in alpine tundra. *Frontiers in Microbiology*, **6**, doi: 10.3389/fmicb.2015.00582.

Shi, S., Nuccio, E., Herman, D. J., Rijkers, R., Estera, K., Li, J., da Rocha, U. N., He, Z., Pett-Ridge, J., Brodie, E. L., Zhou, J., Firestone, M. 2015. Successional Trajectories of Rhizosphere Bacterial Communities over Consecutive Seasons. *MBio*, **6**, doi: 10.1128/mBio.00746-15.

Singh, R. P., Shelke, G. M., Kumar, A. & Jha, P. N. 2015. Biochemistry and genetics of ACC deaminase: a weapon to "stress ethylene" produced in plants. *Frontiers in Microbiology*, **6**, doi: 10.3389/fmicb.2015.00937.

Sirvent, T. & Gibson, D. 2002. Induction of hypericins and hyperforin in *Hypericum perforatum* L. in response to biotic and chemical elicitors. *Physiological and Molecular Plant Pathology*, **60**, 311-320, doi: 10.1006/pmpp.2002.0410.

Sirvent, T. M., Krasnoff, S. B. & Gibson, D. M. 2003. Induction of hypericins and hyperforins in *Hypericum perforatum* in response to damage by herbivores. *Journal of Chemical Ecology*, **29**, 2667-2681, doi: 10.1023/b:joec.0000008011.77213.64.

Sly, L. I., Hodgkinson, M. C. & Arunpairojana, V. 1988. Effect of water velocity on the early development of Mn depositing biofilm in a drinking water system. *FEMS Microbiology Ecology*, **53**, 175-186, doi: 10.1016/0378-1097(88)90440-5.

Smelcerovic, A. & Spiteller, M. 2006. Phytochemical analysis of nine Hypericum L. species from Serbia and the F.Y.R. Macedonia. *Pharmazie*, **61**, 251-252.

Stojanović, G., Đorđević, A. & Šmelcerović, A. 2013. Do other Hypericum species have medical potential as St. John's wort (Hypericum perforatum)? *Current Medicinal Chemistry*, **20**, 2273-2295, doi: 10.2174/0929867311320180001.

Suárez-Moreno, Z. R., Vinchira-Villarraga, D. M., Vergara-Morales, D. I., Castellanos, L., Ramos, F. A., Guarnaccia, C., Degrassi, G., Venturi, V. & Moreno-Sarmiento, N. 2019. Plant-Growth Promotion and Biocontrol Properties of Three *Streptomyces* spp. Isolates to Control Bacterial Rice Pathogens. *Frontiers in Microbiology*, **10**, doi: 10.3389/fmicb.2019.00290.

van Passel, M. W. J, Kant, R., Palva, A., Copeland, A., Lucas, S., Lapidus, A., del Rio, T. G. Pitluck, S., Goltsman, E., Clum, A., Sun, H., Schmutz, J., Larimer, F. W., Land, M. L., Hauser, L., Kyrpides, N., Mikhailova, N., Richardson, P. P., Janssen, P. H. de Vos, W. H. & Smidt, H. 2011. Genome sequence

of the Verrucomicrobium *Opitutus terrae* PB90-1, an abundant inhabitant of rice paddy soil ecosystems. *Journal of Bacteriology*, **193**, 2367–2368, doi: 10.1128/JB.00228-11.

Verma, J. P., Yadav, J., Tiwari, K. N. & Kumar, A. 2013. Effect of indigenous *Mesorhizobium* spp. and plant growth promoting rhizobacteria on yields and nutrients uptake of chickpea (*Cicer arietinum* L.) under sustainable agriculture. *Ecological Engineering*, **51**, 282-286, doi: 10.1016/j.ecoleng.2012.12.022.

Vieira, S., Sikorski, J., Gebala, A., Boeddinghaus, R. S., Marhan, S., Rennert, T., Kandeler, E. & Overmann, J. 2020a. Bacterial colonization of minerals in grassland soils is selective and highly dynamic. *Environmental Microbiology*, **22**, 917-933, doi:10.1111/1462-2920.14751.

Vieira, S., Sikorski, J., Dietz, S. Herz, K., Schrumpf, M., Bruelheide, H., Scheel, D., Friedrich, M. W. & Overmann, J. 2020b. Drivers of the composition of active rhizosphere bacterial communities in temperate grasslands. *The ISME Journal*, **14,** 463-475, doi: 10.1038/s41396-019-0543-4.

Wang, G., Mayes, M. A., Gu, L. & Schadt, C. W. 2014. Representation of Dormant and Active Microbial Dynamics for Ecosystem Modeling. *PLoS ONE,* **9**, doi: 10.1371/journal.pone.0089252.

Wang, L., Yang, F. E. Y., Yuan, J., Raza, W., Huang, Q. & Shen, Q. 2016. Long-Term Application of Bioorganic Fertilizers Improved Soil Biochemical Properties and Microbial Communities of an Apple Orchard Soil. *Frontiers in Microbiology*, **7**, doi: 10.3389/fmicb.2016.01893.

Wang, Z., Cao, Y., Zhu-Barker, X., Nicol, G. W., Wright, A. L., Jia, Z. & Jiang, X. 2019. Comammox *Nitrospira* clade B contributes to nitrification in soil. *Soil Biology and Biochemistry*, **135**, 392-395, doi: 10.1016/j.soilbio.2019.06.004.

Wang, Z., Zhang, J., Wu, F. & Zhou, X. Changes in rhizosphere microbial communities in potted cucumber seedlings treated with syringic acid. *PLoS ONE*, **13**, doi: 10.1371/journal.pone.0200007.

Wei, H. -L. & Zhang, L. -Q. 2006. Quorum-sensing system influences root colonization and biological control ability in *Pseudomonas fluorescens* 2P24*. Antonie Van Leeuwenhoek*, **89**, 267-280, doi: 10.1007/s10482-005-9028-8.

Whitham, T. G., Young, W. P., Martinsen, G. D., Gehring, C. A., Schweitzer, J. A., Shuster, S. M., Wimp, G. M., Fischer, D. G., Bailey, J. K., Lindroth, R. L., Woolbright, S. & Kuske, C. R. 2003. Community and ecosystem genetics: a consequence of the extended phenotype. *Ecology*, **84**, 559-573, doi: 10.1890/0012-9658(2003)084[0559:CAEGAC]2.0.CO;2.

Wu, Q., Xiao, J., Fu, L., Ma, M. & Peng, S. 2020. Microporous intermittent aeration vertical flow constructed wetlands for eutrophic water improvement. *Environmental Science and Pollution Research*, **27**, 16574-16583, doi: 10.1007/s11356-020-08067-x.

Xie, C. -H. & Yokota, A. 2005. Reclassification of *Alcaligenes latus* strains IAM 12599T and IAM 12664 and *Pseudomonas saccharophila* as *Azohydromonas lata* gen. nov., comb. nov*., Azohydromonas australica* sp. nov. and *Pelomonas saccharophila* gen. nov., comb. nov., respectively. *International Journal of Systematic and Evolutionary Microbiology*, **55**, 2419-2425, doi: 10.1099/ijs.0.63733-0.

Xu, Z., Zhang, Y., Muhammad, Y. & Wang, G. 2020. *Sphingomonas montanisoli* sp. nov., isolated from mountain soil. **International Journal of Systematic and Evolutionary Microbiology**, **70**, 3606-3613, doi: 10.1099/ijsem.0.004187.

Yadav, A. N., Sachan, S. G., Verma, P. & Saxena, A. K. 2015. Prospecting cold deserts of north western Himalayas for microbial diversity and plant growth promoting attributes. *Journal of Bioscience and Bioengineering*, **119**, 683-693, doi: 10.1016/j.jbiosc.2014.11.006.

Yan, Z. -F., Lin, P., Wang, Y. -S., Gao, W., Li, C. -T., Kook, M. -C. & Yi, T. -H. 2016. *Niastella hibisci* sp. nov., isolated from rhizosphere soil of mugunghwa, the Korean national flow. **International Journal of Systematic and Evolutionary Microbiology**, **66**, 5218-5222, doi: 10.1099/ijsem.0.001498.

Yang, C., Hamel, C., Gan, Y. & Vujanovic, V. 2013. Pyrosequencing reveals how pulses influence rhizobacterial communities with feedback on wheat growth in the semiarid Prairie. *Plant and Soil*, **367**, 493-505, doi: 10.1007/s11104-012-1485-z.

Yang, Y., Wang, N., Guo, X., Zhang, Y. & Ye, B. 2017. Comparative analysis of bacterial community structure in the rhizosphere of maize by high-throughput pyrosequencing. *PLoS ONE*, **12**, doi: 10.1371/journal.pone.0178425.

You, C., Zhang, C., Kong, F., Feng, C., & Wang, J. 2016. Comparison of the effects of biocontrol agent *Bacillus subtilis* and fungicide metalaxyl–mancozeb on bacterial communities in tobacco rhizospheric soil. *Ecological engineering*, **91**, 119-125. doi: 10.1016/j.ecoleng.2016.02.011.

Yarwood, S., Brewer, E., Yarwood, R., Lajtha, K. & Myrold, D. 2013. Soil microbe active community composition and capability of responding to litter addition after 12 years of no inputs. *Applied and Environmental Microbiology*, **79**, 1385-1392, doi:10.1128/AEM.03181-12.

Yumoto, I., Yamazaki, K., Hishinuma, M., Nodasaka, Y., Suemori, A., Nakajima, K., Inoue, N. & Kawasaki, K. 2001. *Pseudomonas alcaliphila* sp. nov., a novel facultatively psychrophilic alkaliphile isolated from seawater. *International Journal of Systematic and Evolutionary Microbiology*, **51**, 349-355, doi: 10.1099/00207713-51-2-349.

Zamioudis, C. & Pieterse, C. M. 2012. Modulation of host immunity by beneficial microbes. *Molecular Plant-Microbe Interactions*, **25**, 139-150, doi: 10.1094/MPMI-06-11-0179.

Zecchin, S., Corsini, A., Martin, M. & Cavalca, L. 2017. Influence of water management on the active root-associated microbiota involved in arsenic, iron, and sulfur cycles in rice paddies. *Applied Microbiology and Biotechnology*, **101**, 6725-6738, doi: 10.1007/s00253-017-8382-6.

Zhang, L. -N., Wang, D. -C., Hu, Q., Dai, X. -Q., Xie, Y. -S., Li, Q., Liu, H. -M. & Guo, J. -H. 2019. Consortium of Plant Growth-Promoting Rhizobacteria Strains Suppresses Sweet Pepper Disease by Altering the Rhizosphere Microbiota. *Frontiers in Microbiology*, **10**, doi: 10.3389/fmicb.2019.01668

Zhang, X. F., Xu, S. J., Li, C. M., Zhao, L., Feng, H. Y., Yue, G. Y., Ren, Z. W. & Cheng, G. D. 2014.The soil carbon/nitrogen ratio and moisture affect microbial community structures in alkaline permafrost-affected soils with different vegetation types on the Tibetan Plateau. *Research in Microbiology*, **165**, 128-139, doi:10.1016/j.resmic.2014.01.002

Chapter 5 – High throughput cultivation to retrieve isolates with taxonomic novelty

5.1 – Results

5.1.1 – Halle samples that were selected for cultivation

A high throughput cultivation effort was performed shortly after bulk soil and the rhizosphere of *Hypericum* plants were collected from IPB Halle in 2016. The cultivation aims to assist the study of bacterial communities of *Hypericum* plants to gain further insights regarding the link between bacterial communities associated to *Hypericum* and the production of hypericin and hyperforin with anti-depression effect. Thus, our target is to isolate bacterial taxa that may have an impact in the production of the compounds in their host plant. Since phylogenetically novel taxa are likely to harbour functions that are less studied, including production of hypericin and hyperforin in *Hypericum* plants, they were specifically targeted. A total of 6 specific soil samples were selected for high throughput cultivation approach (Table 8) with each representing bulk soil or rhizosphere of wild *H. perforatum* and *H. polyphyllum*. Two rhizosphere samples of *H. perforatum* was selected instead of just one as the plant species is the most well-known producer of hypericin and hyperforin. In addition, rhizosphere samples of greenhouse cultivated *H. perforatum* were available and thus also selected for cultivation.

Table 8 – Soil samples collected from Halle in 2016 that were selected for cultivation.

Sample code	Habitat	Plant species	Origin
2R	Rhizosphere	*H. polyphyllum*	Field no. 2
3BS	Bulk soil	*H. polyphyllum*	Field no. 2
5R	Rhizosphere	*H. perforatum*	Greenhouse
7R	Rhizosphere	*H. perforatum*	Field no. 1
9BS	Bulk soil	*H. perforatum*	Field no.1
9R	Rhizosphere	*H. perforatum*	Field no.1

These samples were also included in the analysis of bacterial communities of *Hypericum* in the first scoping study (see sub-chapter 4.1.1). Rarefaction curves of these 6 samples indicated that our sequence inventory cover most of the taxa present in each sample (Figure 43A) with sample coverage estimates ranged between 0.97% and 0.99%. It was evident that greenhouse samples have the lowest bacterial diversity (may be caused by the heat sterilization process that was applied to the greenhouse soil in prior), as also confirmed by the result of alpha diversity after the samples were normalized to the same number of reads (Figure 43B). Thus, it is possible that we retrieve less diverse bacterial taxa from greenhouse samples.

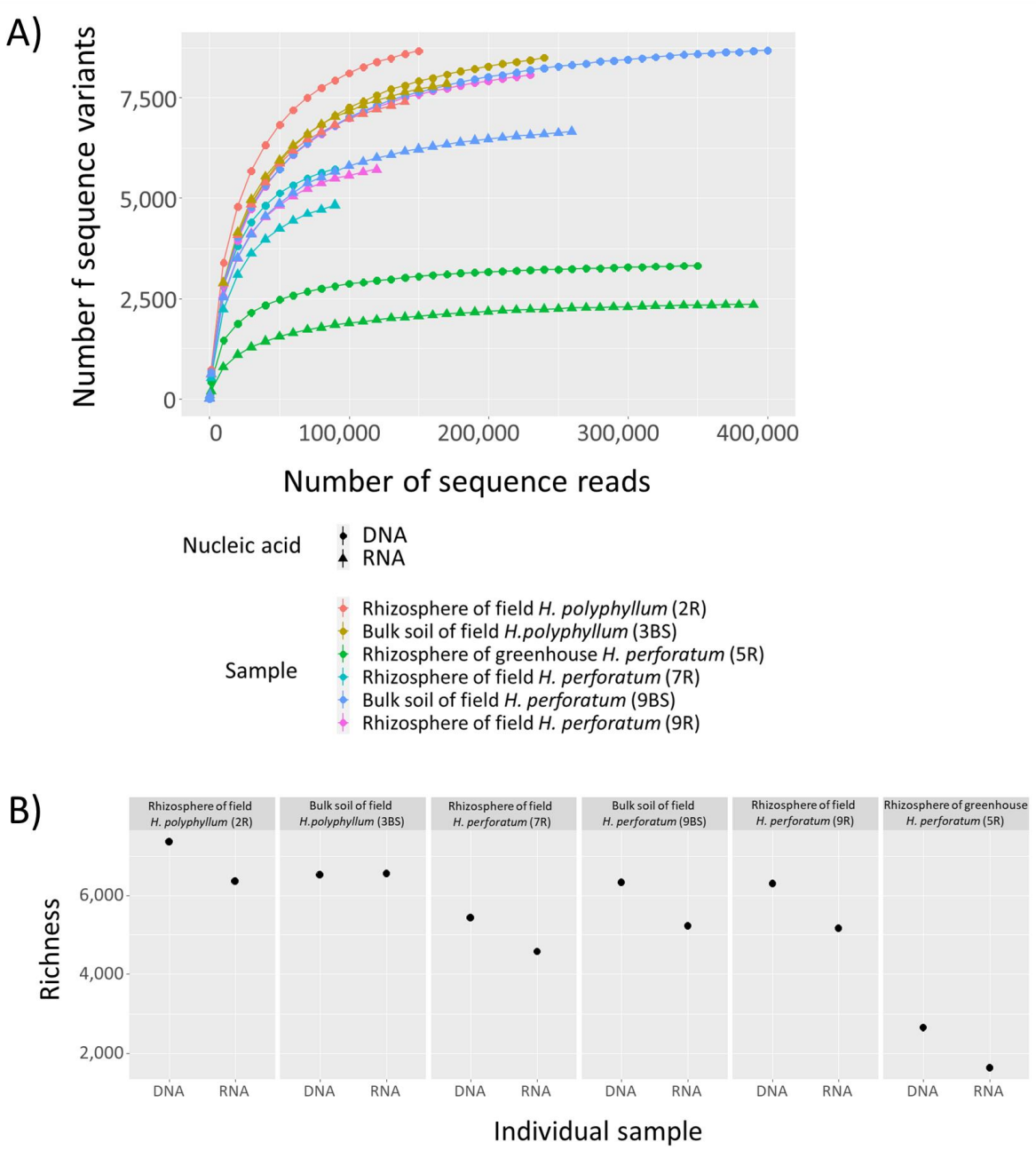


Figure 43 – Rarefaction curves for bulk soil and the rhizosphere of *H. perforatum* and *H. polyphyllum* that were selected for cultivation (A) and the species richness at sequence variant level.

Moreover, NMDS plot based on weighted UniFrac distances revealed separation of greenhouse samples from the rest (Figure 44), implying remarkable differences in bacterial community structure and composition between greenhouse and field samples. The differences are likely due to different soil physicochemical properties across samples (in addition to the impact of heat sterilization process of greenhouse soil). Since bacterial composition of greenhouse samples were quite distinct from others, we may assume that the bacterial taxa isolated from greenhouse samples will be quite different from those of the field samples.

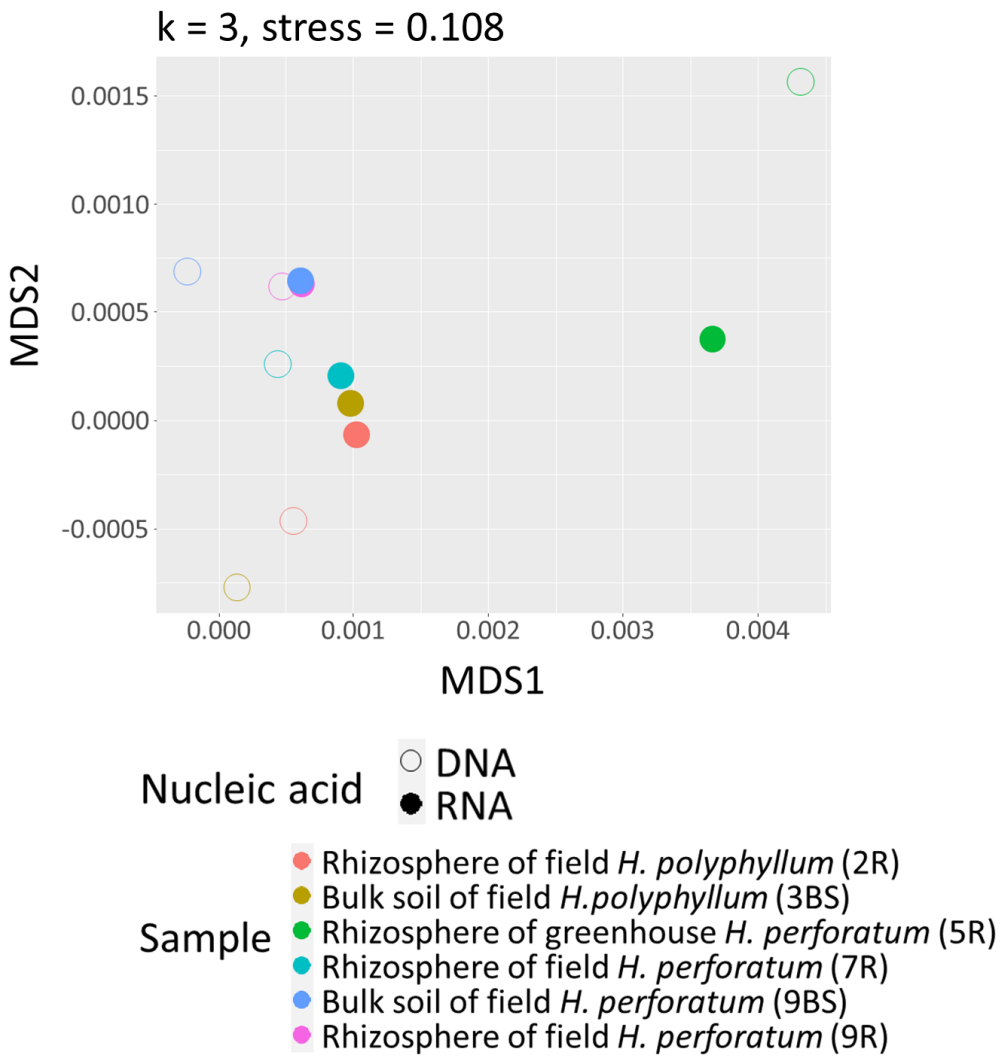
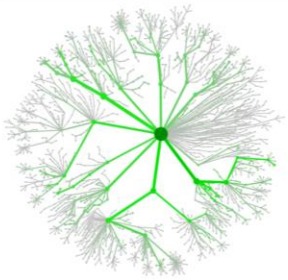
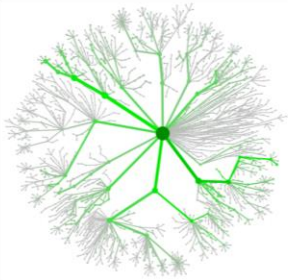
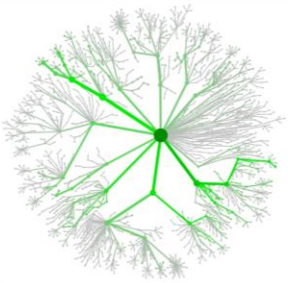
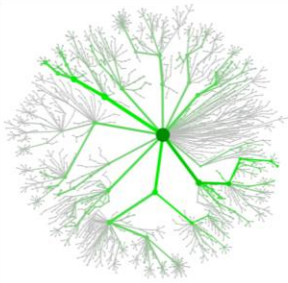
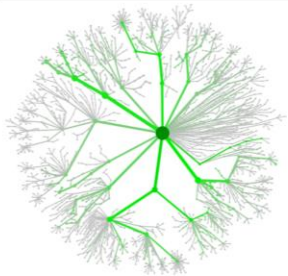
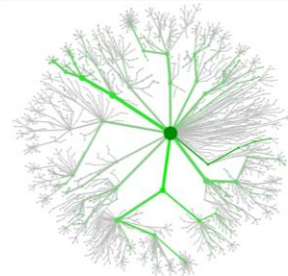
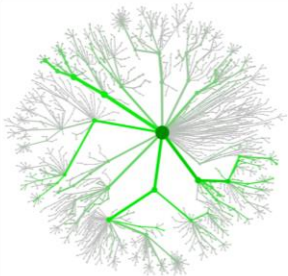
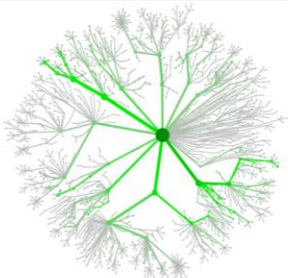
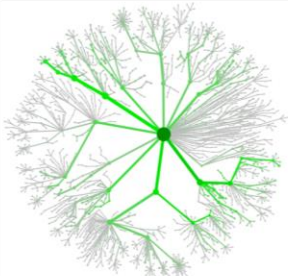
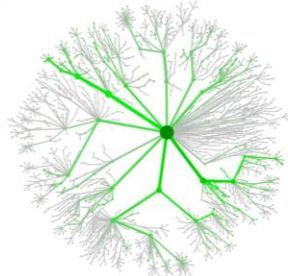


Figure 44 – NMDS plot based on weighted UniFrac distances, comparing the total (DNA-based) and active (RNA-based) bacterial community composition at sequence variant level between soil samples that were selected for cultivation.

Both total (DNA-based) and active (RNA-based) bacterial community composition of each sample was illustrated using the R package metacoder up to genus level (Figure 45). Actinobacteriota, Proteobacteria, Chloroflexi, Gemmatimonadota, Firmicutes, Planctomycetota, Acidobacteriota, Bacteriodota, and Verrucomicrobiota were among the abundant phyla across all samples. At class level, these included *Actinobacteria, Alphaproteobacteria, Gammaproteobacteria, Bacilli,* and *Bacteroidia*. Most of Planctomycetota reads belonged to *Gemmataceae* and *Pirellulaceae*. As previously observed (Figure 14, chapter 4), members of the phylum Cyanobacteria classified as *Nostococaceae,* were dominant in the greenhouse samples, especially in the active bacterial communities (RNA-based; Figure 45).

A)

Sample	DNA	RNA
Rhizosphere of field *H. polyphyllum* (2R)		
Bulk soil of field *H. polyphyllum* (3 BS)		
Rhizosphere of green-house *H. perforatum* (5R)		
Rhizosphere of field *H. perforatum* (7R)		
Bulk soil of field *H. perforatum* (9BS)		
Rhizosphere of field *H. perforatum* (9R)		

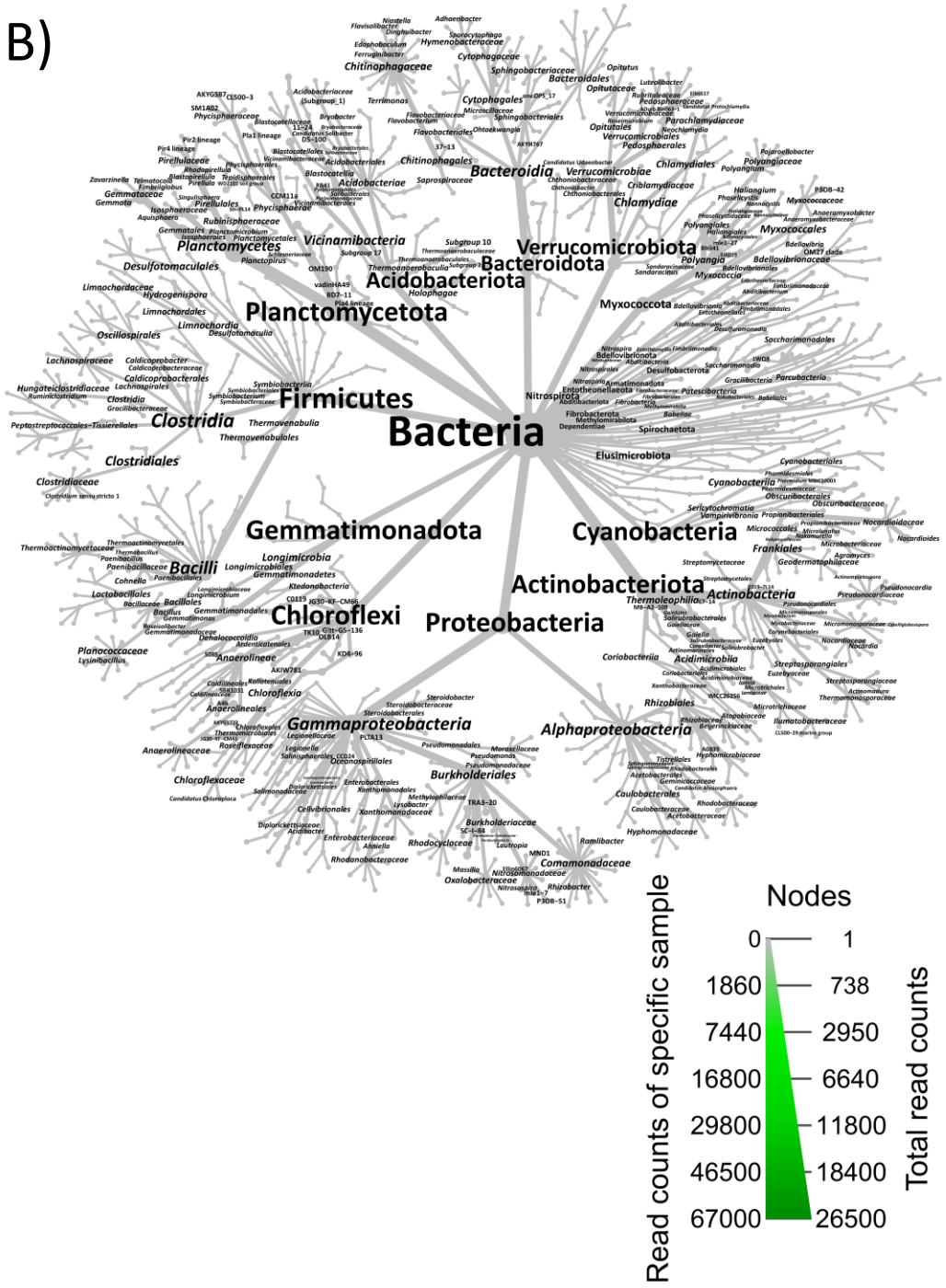


Figure 45 – Bacterial community composition of samples selected for cultivation. Highlighted with green color are the taxa identified for each sample and color intensity reflects the number of taxa that were classified as that taxon (the values are given in Figure B). Figure B depicts the key tree with the taxonomic information. Node diameter reflects number of sequence variant classified as that taxon whereas edge width reflects number of reads.

5.1.2 – High throughput cultivation with employment of growth media with low nutrient content

For each sample that was selected for cultivation, approximately 25 or 50 cells were transferred to a 96 well microtiter plate containing either SSE/HD1:10 or R2A1:10 medium. The inventory of each well was investigated through sequencing of the V1-V2 region of 16S rRNA gene sequences. After 3 months incubation, the plates with inoculation of 50 cells were observed to be overgrown and thus excluded from the analysis.

Percentage of grown wells in the microtiter plates was observed to be higher for SSE/HD1:10 medium compared to R2A1:10, except for greenhouse samples (Figure 46). The result may reflect a better representation of soil nutrient content on SSE/HD1:10 medium compared to R2A1:10 and thus more wells were grown when inoculated with the soil samples. However, since percentage of grown wells between the two media is similar for greenhouse samples, the benefit of utilizing SSE/HD1:10 medium to retrieve soil bacteria may vary depending on the soil type.

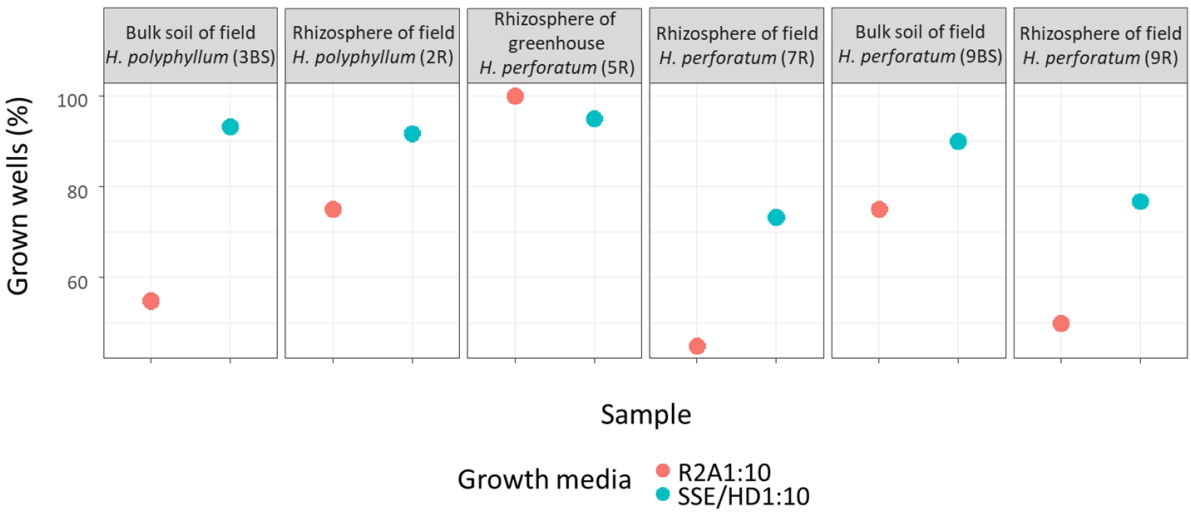


Figure 46 – Percentage of grown wells in 96-well microtiter plates between R2A1:10 and SSE/HD 1:10 media used for cultivation for each soil sample.

After processing of raw reads and removal of sequences associated to chloroplast, a total of 13,429,526 reads belonging to 1727 sequence variants were retrieved. Rarefaction curves at sequence variant level indicated that our sequence inventory covered most of the taxa present in each well (Figure 47A). The data was further rarified (normalized) to 11,102 reads per sample to allow fair comparison between different growth media and source of isolation, leaving the dataset with 1717 sequence variants distributed across grown wells of microtiter plates. Diversity measures including richness, Shannon, and evenness were significantly higher in SSE/HD1:10 medium compared to R2A1:10 (Figure 47B), implying that the former is more suitable to capture soil bacterial diversity.

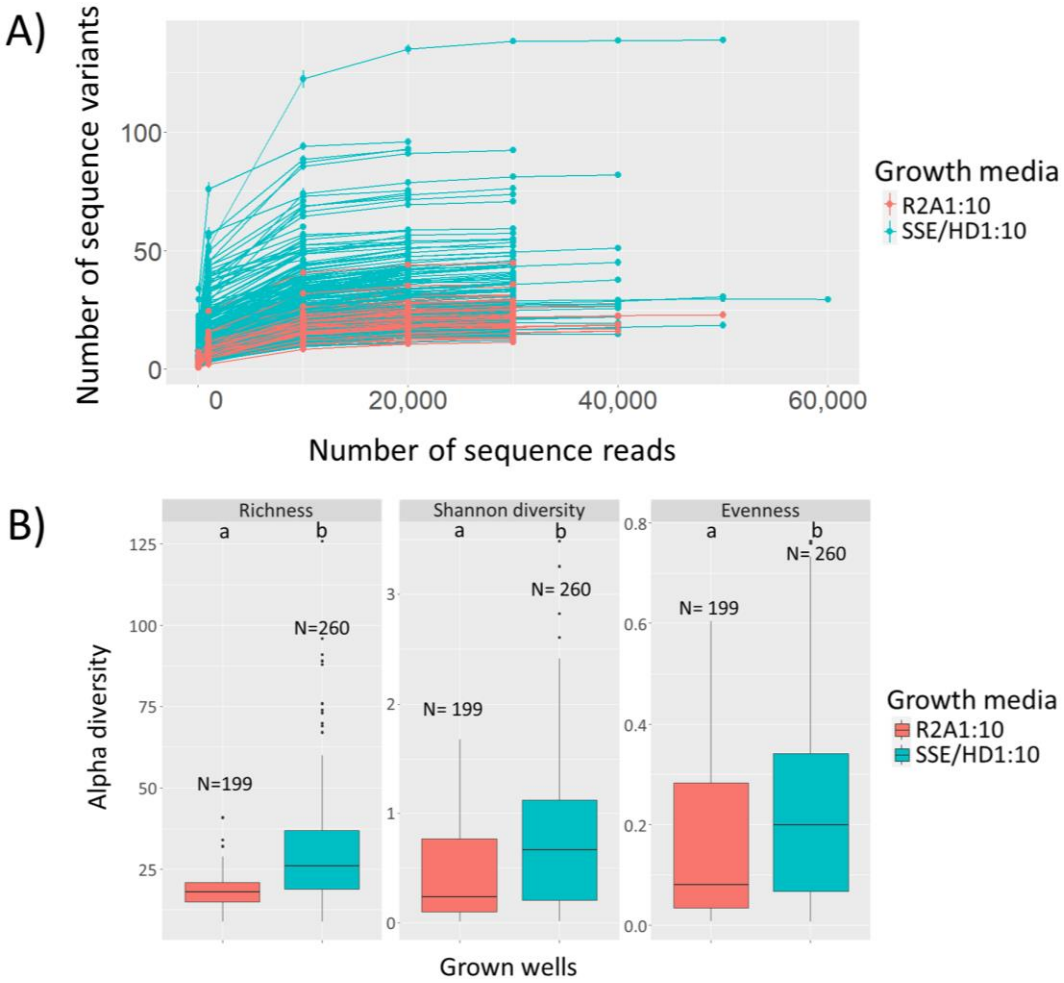


Figure 47 – Rarefaction curves for sequence variants detected in each well of the microtiter plates for both growth media. B) The diversity measures on R2A1:10 and SSE/HD 1:10 media. Letters denote significant differences (t test, $p < 0.05$). N represents the number of samples for each group.

Due to the fact that the greenhouse samples had lower diversity when compared to those of the fields (Figure 43), it is likely that more bacterial taxa will be retrieved from the field samples. However, this was not proven since the wells inoculated with greenhouse samples were observed with higher bacterial richness and Shannon diversity when compared to when they were inoculated with the field samples of *H. perforatum*, especially when SSE/HD1:10 was utilized as the cultivation medium (Figure 48). Thus, it seems that the cultivation media have stronger impact on the bacterial diversity inside the wells compared to the source of isolation.

In order to evaluate further if either SSE/HD1:10 or R2A1:10 preferentially support the growth of only few bacterial taxa, the alpha gambin was calculated for the wells belonging to the same soil source and cultivation medium (Figure 48). The insights can be used in the future, for example if we aim to isolate specific bacterial taxa that are easily overgrown by many other taxa. The growth medium that specifically supports the growth of the target bacteria while inhibit others can be

utilized for the subsequent cultivation approaches targeting bacterial taxa with similar traits. Low alpha gambin values suggest logseries-like species distribution (SAD) where the community only belong to a few species and each species is represented by very few individuals while high values suggest lognormal-like SAD where only few species have either high or low abundances while most have intermediate abundance. The result implies that although the bacterial diversity is lower in R2A1:10 medium when compared to SSE/HD1:10, it does not necessarily mean that the former support the growth of only few taxa, since the alpha gambin values were not significantly different between those two-cultivation media.

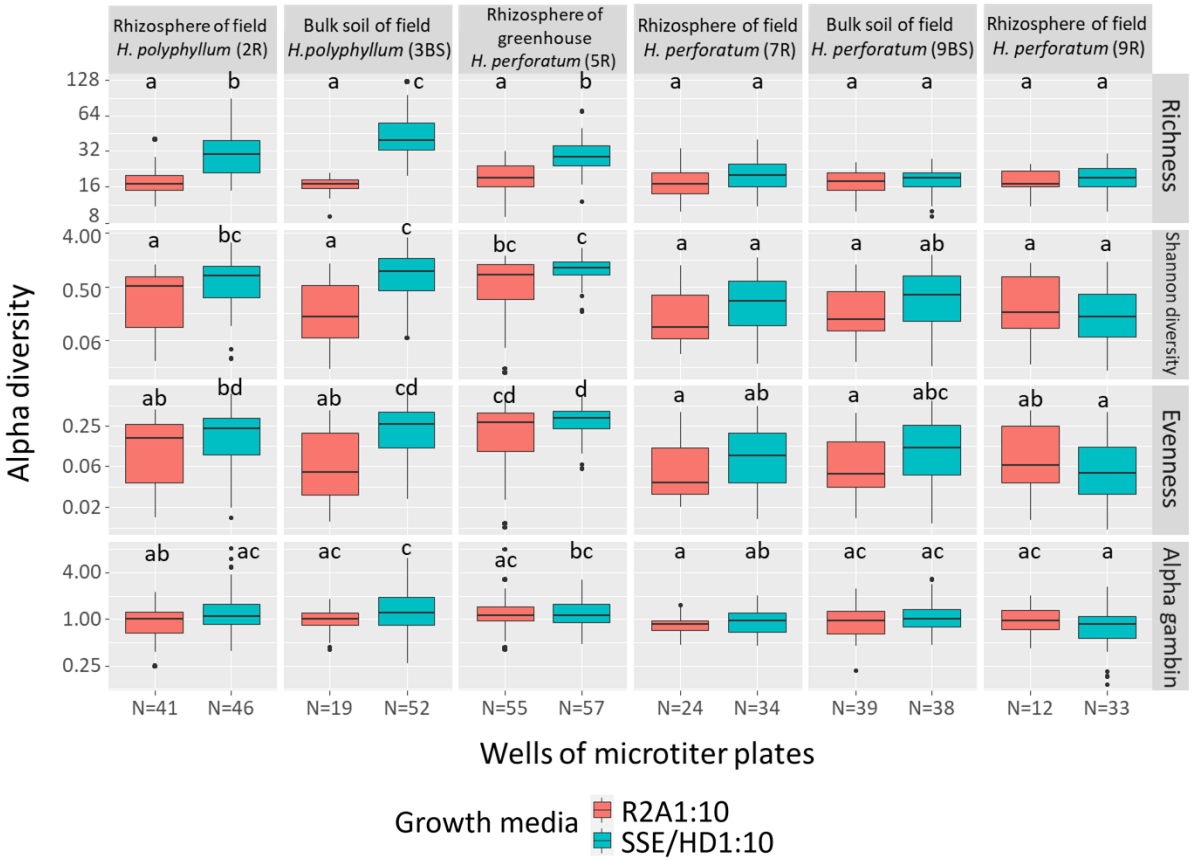


Figure 48 – Diversity measures (richness, Shannon diversity, evenness, and alpha gambin) of grown wells inoculated with 6 different soil samples. Letters at the top denote significant differences between distinct cultivation media (R2A1:10 and SSE/HD1:10) and soil samples (multcomp, $p < 0.05$). N represents the number of grown wells on the microtiter plates.

Due to differences in the region of 16S rRNA gene sequences that was amplified, the comparison between the entire taxa that was present in the soil and the subset that was grown in cultivation media could not be done directly at sequence variant level. Thus, the comparison was performed at genus level, after sequence variants on both datasets were individually assigned using the silva 138 database. Only taxa with abundance value of above 1%, (implying that the taxa are growing) were

included in the analysis. The analysis revealed that for each soil sample, approximately 18 genera in the cultivation media were not actually detected on the original soil samples (employing dataset that has not been rarified) (Figure 49A), including members of *Aminobacter, Novosphingobium, Cutibacterium, Lysinimonas,* and *Nesterenkonia* (Figure 49B). The abundance of these taxa on the soil at the time of the sampling is likely very low and below the threshold for the detection limit, and thus not observed as members of soil or plant microbiomes.

However, the bacterial genera that were detected in both original soil samples and cultivation media are not necessarily abundant either. For example, in the case of rhizosphere of *H. polyphyllum*, even though genera such as *Nocardioides* and *Bacillus* that grew in the cultivation media were abundant in the original soil sample (4.7% and 2.1% relative abundance, respectively) many other genera including *Ohtaekwangia* and *Patulibacter* had very low abundance (0.05% and 0.005%, respectively). Thus, other factors including the employed growth media are likely having more impact during cultivation.

To determine bacterial taxa that were enriched in one of the liquid cultivation media when compared to the other, pairwise comparison of bacteria growing in each medium was performed up to genus level. The result revealed that a higher number of taxa was significantly enriched in SSE/HD1:10 medium when compared to R2A1:10 (Figure 50). The enriched taxa in SSE/HD1:10 medium included members of *Arenimonas, Cryobacterium*, *Cutibacterium, Erythrobacter, Longispora, Lysobacter,* and *Turneriella* whereas *Mycobacterium* was enriched in R2A1:10 medium. Again, this suggests that SSE/HD1:10 medium is more suitable to capture soil bacterial diversity when compared to R2A:10. However, the selection of cultivation media is also depending on the target, since R2A1:10 is a better medium to capture taxa such as members of *Corynebacteriales.*

In order to investigate if some bacterial genera are consistently isolated together in each cultivation media, co-occurrence pattern was determined through calculation of partial correlation coefficients (using R netassoc package) across bacterial taxa to find the pair that are more likely occur together than expected under the null model. The association of distinct bacterial genera, independently of cultivation media, was calculated based on relative abundance data and compared to null model of community assembly (generated with R vegan package). This allows us to study specific interactions among bacterial taxa and the knowledge may be implemented in the future to improve cultivation success. Since bacteria with less than 1% abundance are likely unable to grow on the cultivation media, they were excluded from the analysis. The result revealed 466 positive associations where a total of 39 bacterial genera were found to be positively associated with at least two other genera (Figure 51; $p < 0.01$), regardless of the cultivation media. However, none of the negative associations

across bacterial genera were found to be significant (p value > 0.01). The bacterial genera that co-occur the most with the other genera included *Pseudomonas, Nocardioides, Sphingomonas, Mesorhizobium* and *Reyranella*, with each positively associated with other 27, 27, 26, 25, and 25 other genera, respectively. This may reflect the ability of these genera to promote the growth of other taxa for example through production of some specific metabolites.

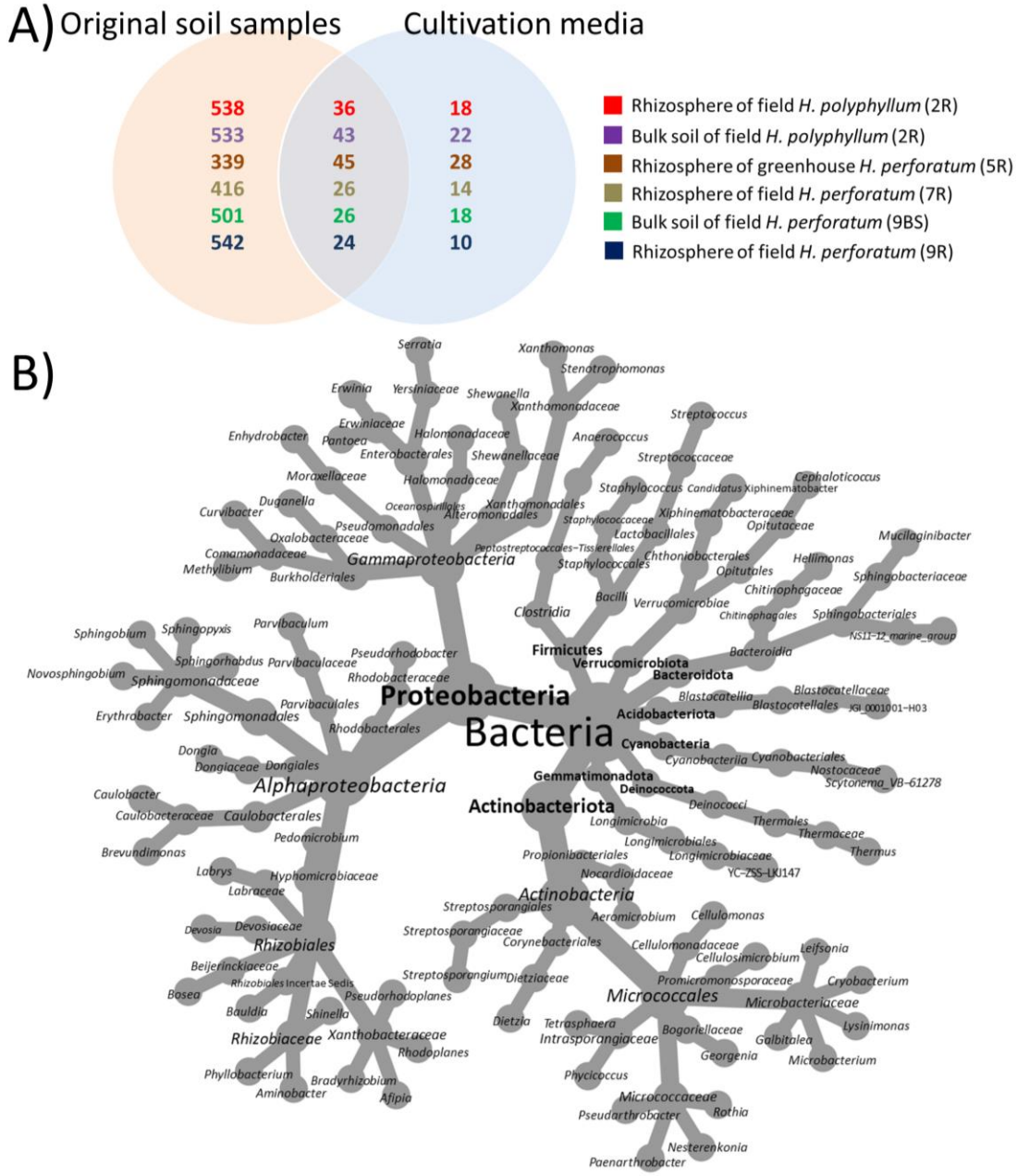


Figure 49 – Unique and overlap taxa at genus level between original soil samples and the subset that were detected in cultivation media for each soil sample that was selected for cultivation (A; grown wells on the microtiter plates belonging to the same soil and cultivation medium were pooled together). B) Unique taxa that were detected only in the cultivation approaches, up to genus level.

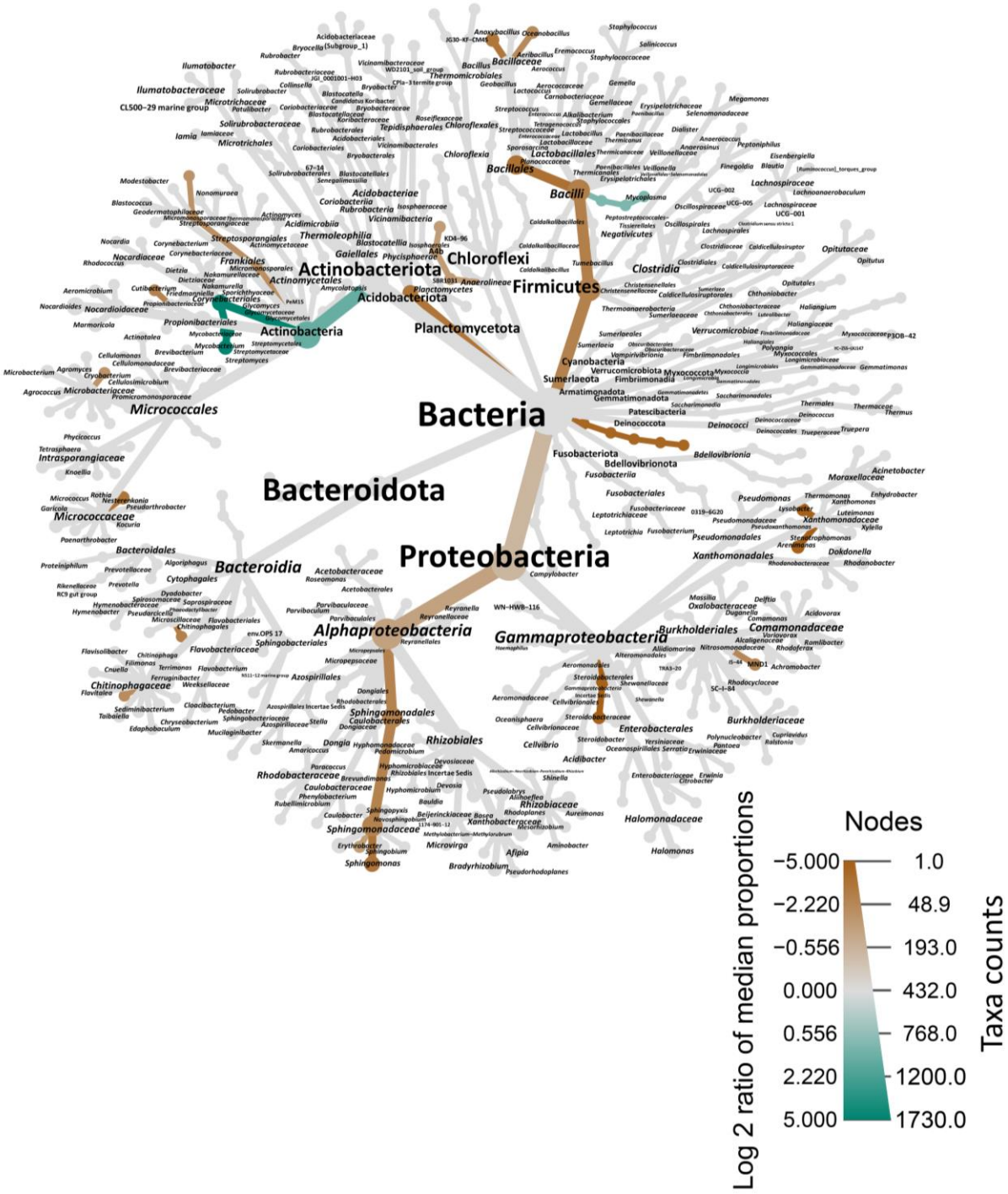


Figure 50 – Pairwise comparison of bacteria growing in liquid SSE/HD 1:10 and R2A 1:10 media, from phylum up to genus level. Brown and green colors represent the significantly enriched taxa in SSE/HD1:10 and R2A 1:10 medium, respectively (Wilcoxon rank-sum test p value < 0.05) and color intensity reflects log2 ratio of median proportions. Node diameter reflects number of taxa classified as that taxon whereas edge width reflects number of reads.

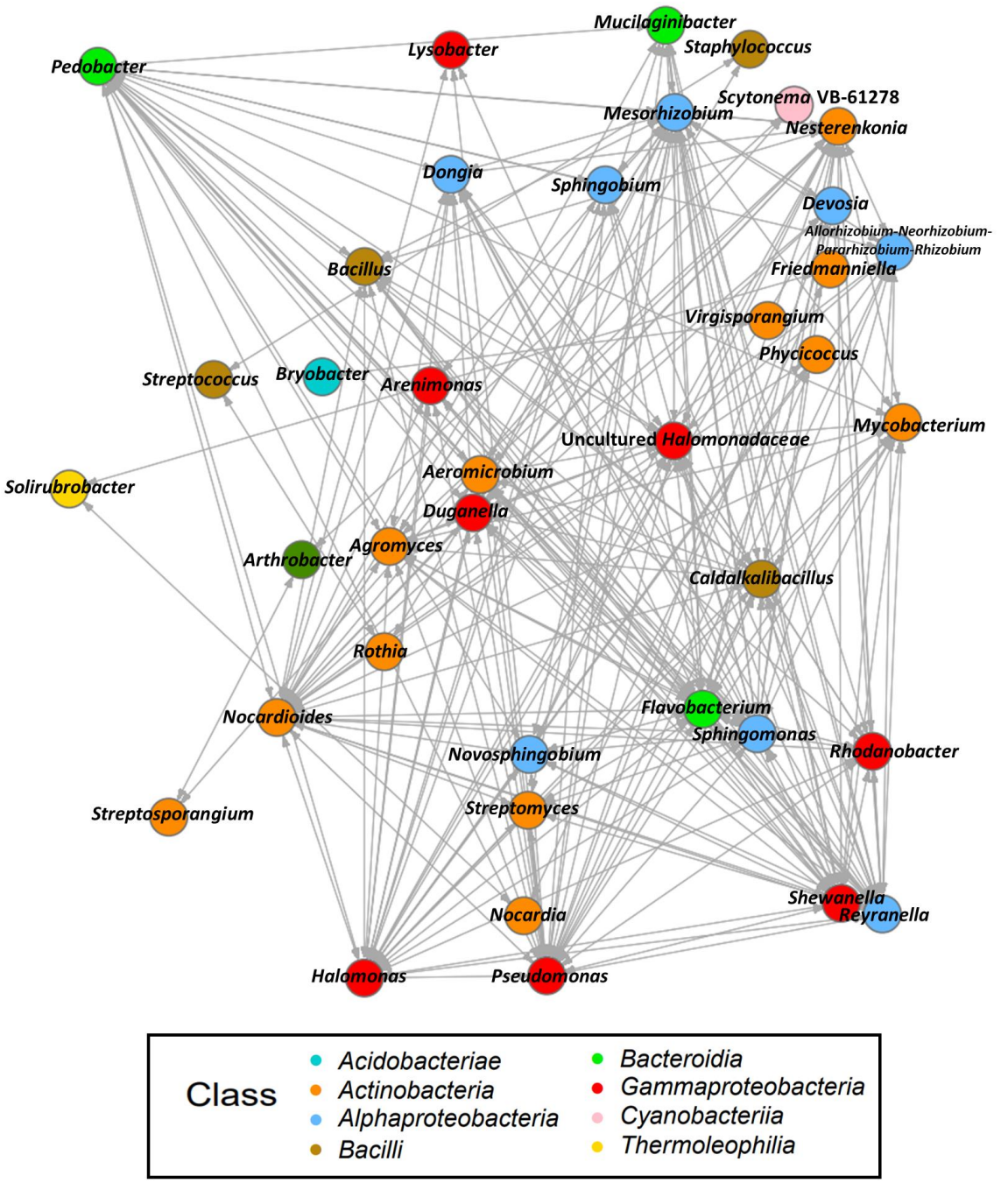


Figure 51 – Significant positive associations ($p < 0.01$) across bacterial genera growing inside the wells of microtiter plates, regardless of the cultivation media. Only those that were positively associated with at least two other bacterial genera are depicted in the plot. Species association was inferred based on the relative abundance data. The partial correlation coefficients were calculated across distinct taxa and compared to null model to determine the pair of taxa that are more likely co-occur together than expected under the null model. P-values were adjusted with the Benjamini-Hochberg (FDR) correction for multiple comparisons. Only taxa with more than 1% abundance (most likely are growing in the cultivation media) were included in the analysis. The taxa were colored according to the bacterial class that they belong to.

5.1.3 – Co-cultivation of a novel member of Acidobacteriota with helper bacteria

Wells with bacterial growth were screened for novel bacterial taxa (less than 97% similarity of 16S rRNA gene sequences with known isolates), since they may represent novel functions that may be related to hypericin and hyperforin production in *Hypericum* plants. Purification by streaking on agar media was employed for the selected wells and results in cultivation of a total of 147 bacterial isolates belonging to 20 different bacterial order, according to the alignment of the partial length of 16S rRNA (~700 bp) to the Eztaxon database (Figure 52). The full-length of 16S rRNA of bacterial isolates (~1400 bp) was used subsequently for the alignment (Table 9). Some novel bacterial taxa that were successfully isolated included members of family *Rhodospirillaceae* (strain R5959 & R5913), *Patulibacteraceae* (strain 8_15), and *Iamiaceae* (strain 8_14).

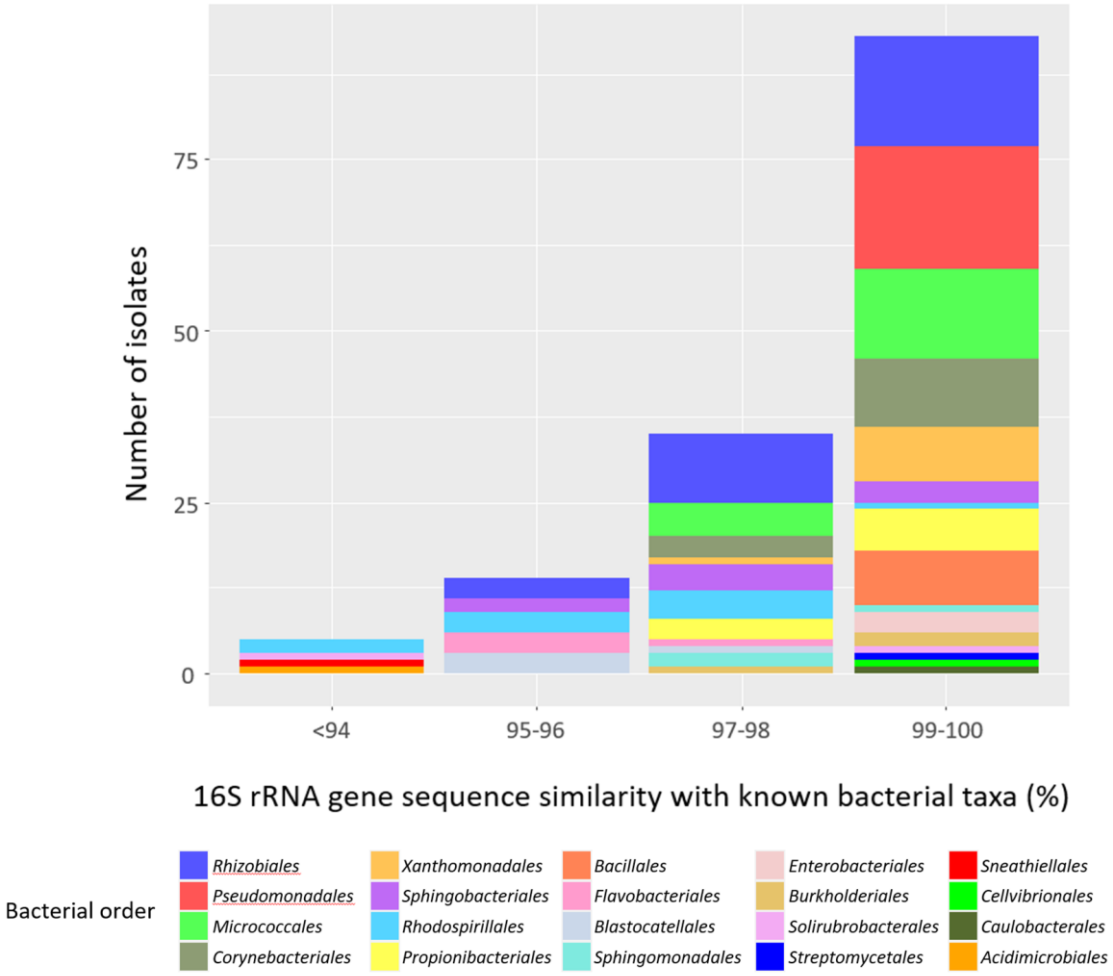


Figure 52 – The similarity of the 16S rRNA gene sequences (partial length; ~700 bp) of the 147 bacterial isolates with known bacterial taxa. The bacterial isolates were colored according to the bacterial order that they belong to.

Table 9 – Isolates in pure culture with less than 97% similarity of the almost full-length 16S rRNA gene sequences (~1400 bp) with known bacterial taxa, and thus may represent novel species (< 97% similarity) or genus (< 95% similarity).

No.	Strain Code	Closest taxon	Similarity
1	R5959	*Dongia mobilis*	91.49%
2	R5913	*Dongia soli*	91.84%
3	8.15	*Patulibacter brassicae*	90.96%
4	8.14	*Aquihabitans daechungensis*	93.44%
5	13.8	*Pseudolabrys taiwanensis*	94.35%
6	10.31	*Flavobacterium paronense*	94.58%
7	9.60	*Flavobacterium chungangense*	95.04%
8	7.4	*Taibaiella soli*	95.68%
9	8.34	*Dongia mobilis*	96.12%
10	9.65	*Oligotropha carboxidovorans*	96.64%
11	6.10	*Stenotrophobacter roseus*	96.65%
12	13.52	*Nocardioides terrigena*	96.91%

Among the cultivated bacterial taxa included one member of Acidobacteriota (6.10) that grew only in close proximity with helper bacteria. As a step closer in understanding the interaction between distinct bacterial taxa, which may improve cultivation success in the future, the effect of helper bacteria to the growth of the Acidobacteriota was further investigated. The bacterium was classified as a member of the genus *Stenotrophobacter* and was first isolated when it was streaked on the agar media in close proximity with *Pseudomonas prosekii* (Figure 53).

Although not shown, other helper bacteria including *P. fluorescens, P. arsenicoxydans, P. lini, P. kilonensis, B. megaterium* and *M. qingshenggi* were also able to promote growth of the target strain. Almost no growth or no growth at all was observed when the strain was grown alongside *P. veronii, P. marginalis,* and *P. panacis.* The diversity of helper bacteria indicated that the requirements to promote growth of target strain is not depending on phylogenetic affiliation. It seems that rather diverse taxa produced specific compounds that are vital for the growth of the Acidobacteriota.

However, the growth of the Acidobacteriota was not positively correlated with the presence of any particular bacterium in the same well of microtiter plate based on the co-occurrence analysis (Figure 51). This occurred due to the low abundance of the Acidobacteriota, which was only around 0.1%, and hence excluded from the network analysis.

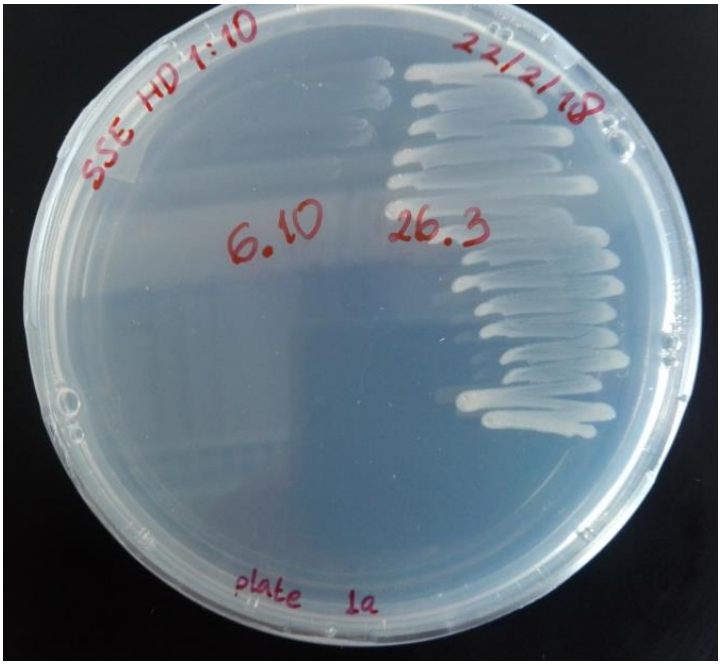


Figure 53 – Co-cultivation of a member of Acidobacteriota (code: 6.10) with *Pseudomonas prosekii* (code: 26.3) in SSE/HD1:10 agar where the growth of the Acidobacteriota was less pronounced as the distance with *P. prosekii* increases.

Furthermore, the Acidobacteriota was grown in liquid SSE/HD1:10 medium with and without the addition of *P. prosekii* supernatant from LB medium, to determine the importance of the helper bacterium for the strain growth. The SSE/HD1:10 and LB media were both employed since the target bacterium preferred the former while the helper bacterium preferred the latter. The result revealed that the growth of the Acidobacteriota was promoted with the addition of the helper bacterium supernatant (Figure 54), which was likely to contain crucial compounds that are vital for the Acidobacteriota growth. The growth of the Acidobacteriota was even better when living cells of helper bacterium was inoculated into the medium instead of only the supernatant. This implies that the helper bacterium actively produces/releases specific compounds that promote the growth of the target bacterium.

A more advanced and thorough study needs to be performed in the future to explore specific interactions between the target and helper bacterium and to reveal the specific compounds that are vital for the growth of the Acidobacteriota. As the majority of bacteria still escape cultivation so far, the study will be advantageous to extend our knowledge regarding the uncultured bacteria and to later improve cultivation success. In the case of *Hypericum* plants microbiome, it may lead to the isolation of novel taxa with novel functions that are related to the plant ecological functions, including production of hypericin and hyperforin.

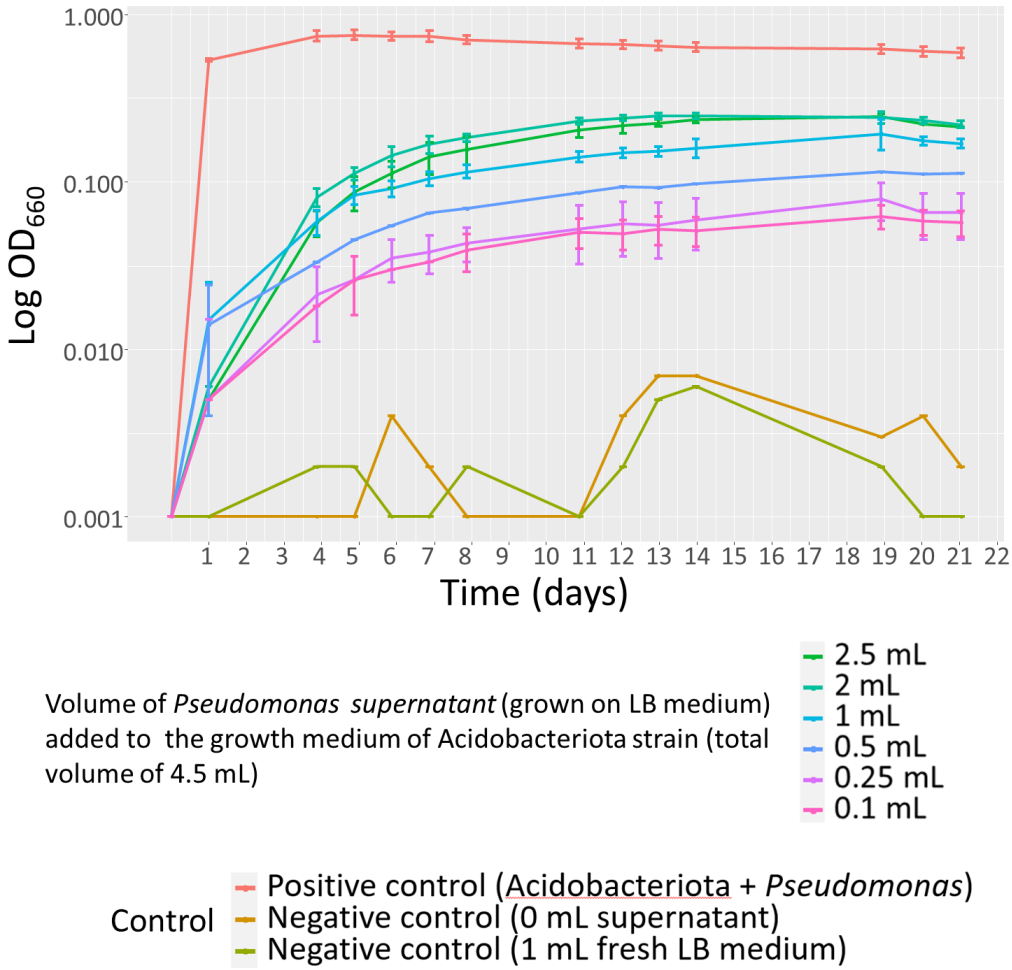


Figure 54 – Growth curves of the novel member of Acidobacteriota on liquid SSE/HD1:10 with the addition of the supernatant of helper bacterium *Pseudomonas prosekii* grown in LB. LB was chosen as the helper growth better in the medium. Positive control was when the Acidobacteriota and the helper bacterium were grown together in SSE/HD1:10. Negative controls only included the Acidobacteriota, either with or without the addition of fresh LB medium, to make sure that there was no effect associated with the growth medium.

5.1.4 – Characterization of two new members of the family *Rhodospirillaceae*

Two members of a novel genus of the family *Rhodospirillaceae*, namely strains R5913^T and R5959^T, were selected among the phylogenetically novel isolates (Table 9) for further characterization based on a higher taxonomic novelty (novel genus instead of species) and readiness to be cultured in the laboratory. The valid description of the strains has been recently published (Noviana *et al.* 2020). The growth of other novel genera still needs to be optimized, for example strain 8.15 that only observed to grow after 3-4 months incubation. The strains were isolated from the rhizosphere of *H. perforatum* grown in the greenhouse of the Leibniz Institute of Plant Biochemistry in Halle (Table 8; sample code 5R). The abundance of the strains was assessed by aligning the sequences of both

strains with all sequences detected in the original soil samples. Both strains had an abundance of approximately 0.02% in their original habitat. They can be considered as rare (low abundant) taxa in the soil sample, where around 10% of all detected taxa had higher abundance compared to them. The percentage of taxa with higher abundance may be small but soil is considered a habitat with a large fraction of rare biosphere (often defined with abundance < 0.01%). The strains were only detected in DNA data and not in RNA data, implying they were not transcriptionally active and thus may not perform important ecological functions related to the host plant.

In order to determine if the strains have a contribution to soil ecosystem functioning or perform functions related to plant fitness in similar habitat, the sequences were aligned using 99% similarity approach against bacterial sequences belonging to soil-related (e.g. sediment and sludge) and rhizosphere habitat (available from the Integrated Microbial Next Generation Sequencing (IMNGS) depository). The result revealed that both strains are more abundant in other soils, with the relative abundance value can be up to 2% (Figure 55), and thus may perform functions related to soil ecosystem functioning albeit not related to plant fitness. This implies that some specific factors in the original soil sample may hinder the growth of both strains and given favourable condition, they may actively contribute to soil processes.

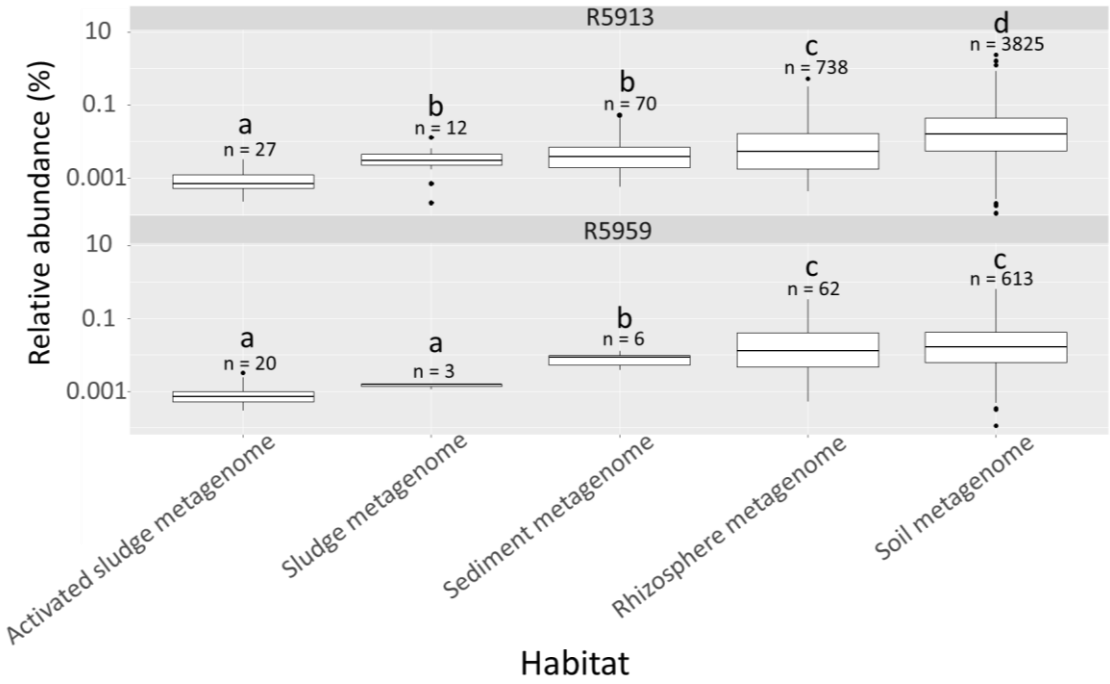


Figure 55 – Relative abundance of strains R5913^T and R5959^T across distinct habitats within the IMNGS depository, aligned at 99% similarity. Letters at the top of each boxplot denote significant differences (multcomp, $p < 0.01$) and the number of datasets where the strains were detected is displayed below each letter.

5.1.4.1 – Chemotaxonomic properties of strains R5913^T and R5959^T

Strains R5913^T and R5959^T were classified as members of *Rhodospirillaceae* and shared 98.6% 16S rRNA gene sequence similarity with one another. According to the EzBioCloud identification tool (Yoon *et al.* 2017a), the closest relatives of strain R5913^T were *Oceanibaculum pacificum* MCCC 1A02656^T (91.9%), *Dongia mobilis* CGMCC 1.7660^T (91.8%) and *Dongia soli* D78^T (91.7%), while the closest relatives of strain R5959^T were *Dongia mobilis* CGMCC 1.7660^T (91.6%), *Oceanibaculum pacificum* MCCC 1A02656^T (91.4%) and *Dongia rigui* 04SU4-P^T (91.3%).

Both strains stained Gram negative and appeared as small, round, smooth, shiny, convex, white colonies with 1 mm in diameter on R2A agar. Some cells were motile by means of a single polar flagellum (Figure 56). Capsules and endospores were not observed. Within their closest relatives, capsules are only present in *Lacibacterium aquatile* LTC-2^T (Table 10). Cells of R5913^T were single-short-rods, or sometimes slightly curved (0.7-1.5 µm long and 0.5-0.8 µm wide), as observed under the light microscope (Zeiss Axio Lab. A1; Carl Zeiss). Cells of strain R5959^T were slightly curved, short, and rod-shaped (0.5-2.5 µm long and 0.3-1.0 µm wide) and occurred as either single cell or in short chains. Rod-shaped cells are a common feature of R5913^T and R5959^T closest relatives in the family *Rhodospirillaceae* (Table 10), with the exception of the star-shaped cells of *Stella* (currently classified as a member of the family *Acetobacteraceae*) (Vasilyeva 1985). Catalase activity was a feature distinguishing the strains strain since R5913^T tested positive and R5959^T negative. Both strains were positive for cytochrome-*c* oxidase which is in agreement with the majority of their closest genera (Table 10).

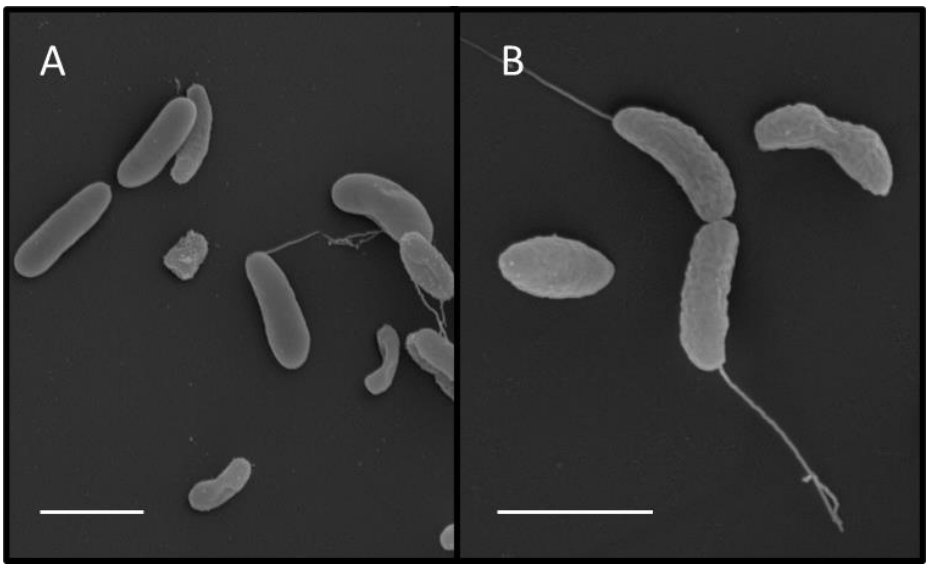


Figure 56 – Scanning electron micrographs of strains R5913^T (a) and R5959^T (b). Scale bars, 1 µm.

Table 10 – Phenotypic characteristics of strains R5913^T and R5959^T compared with other related genera in the family *Rhodospirillaceae.*

Strains/genera: 1, R5913^T; 2, R5959^T; 3, *Reyranella* (Pagnier *et al.* 2011; Kim *et al*. 2013; Lee & Whang 2014; Lee *et al.* 2017; Cui *et al.* 2017); 4, *Tagea* (Jean *et al.* 2016); 5, *Oceanibaculum* (Dong *et al.* 2010; Lai *et al.* 2009; Du *et al.* 2017); 6, *Nisaea* (Urios *et al.* 2008); 7, *Thalassobaculum* (Zhang *et al.* 2008; Urios *et al.* 2010; Su *et al.* 2016); 8, *Dongia* (Kim *et al.* 2016; Liu *et al.* 2010; Baik *et al.* 2013); 9, *Inquilinus* (Jung *et al.* 2011; Coenye *et al.* 2002); 10, *Aliidongia* (Chen *et al.* 2007); 11, *Stella* (Vasilyeva 1985); 12, *Lacibacterium* (Sheu *et al.* 2013) ; 13, *Elstera* (Rahalkar *et al.* 2012; Cai *et al.* 2017); +, positive; -, negative; V, variable response; NA, no data available.

Characteristics	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
Colony pigmentation	White	White	Grey-white, milk colored	Greyish	Grey colored	Cream	Cream yellow, yellowish-brown	White	Creamy white	White	Milky white, greyish white	Creamy white	Milky white to creamed colored, light brown
Cell shape	Short to slightly curved rods	Slightly curved short rods	Rods	Rods	Rods	Pleomor-phic rods	Slightly curved to straight rods	Slightly curved to straight rods	Rods	Rods	Star-shaped	Slightly curved rods	Slightly curved rods
Flagellar motility	+	+	-	+	+	+	V	+	-	+	-	+	V
Capsules	-	-	NA	NA	NA	NA	NA	NA	NA	NA	NA	+	NA
Catalase	+	-	V	+	+	+	+	V	+	+	+	+	+
Oxidase	+	+	+	+	V	+	+	V	+	+	+	+	+
Nitrate reduction	-	-	V	+	V	+	V	V	-	+	NA	-	V
Salt is required for growth	-	-	-	+	-	-	V	-	V	-	-	-	-
Temperature range (°C) (optimum)	20-36 (27-33)	24-40 (33-36)	15-37 (20-35)	20-42 (30-35)	10-45 (25-37)	15-44 (30)	10-40 (30-35)	15-40 (25-37)	15-42	10-37 (28)	(28-30)	10-37 (20-30)	10-37 (20-28)
pH range (optimum)	6.0-8.4 (6.5-7.5)	5.0-8.2 (5.6-7.4)	4.0-10.0 (5.0-9.0)	6.0-9.0 (7.0-8.0)	6.0-11.0 (7.0-9.0)	5.0-9.0 (6.0)	5.0-10.0 (6.5-8.0)	5.0-11.0 (7.0-7.5)	NA	4.5-7.5 (6.0-6.5)	(near neutral/ slightly alkaline)	6.0-9.0 (7.0-8.0)	5.0-9.0 (6.5-7.0)
Glucose fermentation / acidification	-	-	-	-	-	NA	-	-	-	-	NA	+	+
Relation to oxygen †	AE, MAE	AE, MAE	AE, MAE	AE	AE	FA	FA, AE	AE	AE	AE	AE, MAE	FA	AE, MAE
DNA G+C content (% or mol%)	66.02	67.35	59.8-66.5	56.4	64.8-67.7	60.1-60.2	65-69	54.7-71.5	69-9-70.3	65.8	69.3-73.5	58.5	61.2-62.4

Major quinone (s)	Q-10	Q-10	Q-10, Q-9	Q-10	Q-10	Q-10	Q-10	Q-10	Q-10	Q-10	NA	Q-10	Q-10
Major polar lipids:													
Phosphatidylethanolamine	+	+	+	NA	+	-	-	+	NA	+	NA	+	+
Hydroxyphosphotidylethanolamine	-	-	-	NA	-	-	-	+	NA	-	NA	-	-
Phosphatidylmonomethylethanolamine	-	-	+	NA	+	-	-	-	NA	-	NA	+	+
Phosphatidyldimethylethanolamine	-	-	-	NA	-	-	-	-	NA	+	NA	+	-
Phosphatidylglycerol	-	-	+	NA	+	+	+	+	NA	+	NA	+	+
Diphosphatidylglycerol	-	-	+	NA	+	-	-	-	NA	+	NA	-	-
Phosphatidylcholine	-	-	-	NA	-	-	-	-	NA	-	NA	-	-
Major fatty acids													
C19:0 cyclo *ω8c*	+	+	+	+	+	-	+	+	+	+	NA	-	-
C18:1*ω7c*	+	-	+	+	+	+	+	+	+	+	NA	+	+
C16:0	+	+	+	-	+	+	+	+	-	-	NA	+	+
Isolation source	Rhizosphere	Rhizo-sphere	Forest and agricultural soil, bamboo litter, river, lake	Coastal seawater	Surface and deep seawater, hydrothermal vent	Surface seawater	Coastal and deep seawater	Freshwater, activated sludge of a sequencing batch reactor, soil	Respiratory secretions of a CF patient, soil	Forest soil	Soils; freshwater sediments, sewage, sludge, horse manure	Freshwater lake	Stones biofilms, lake

†AE, aerobic; AN, anaerobic; FA, facultative anaerobic; MAE, microaerophilic.

Both strains grew under microaerophilic but not in anaerobic conditions. The results are in accordance with what has been observed for the genus *Stella* (Vasilyeva 1985) and some members of *Reyranella* (Pagnier *et al.* 2011; Kim *et al.* 2013) and *Elstera* (Rahalkar *et al.* 2012) (Table 10). Most members of the family *Rhodospirillaceae* grow aerobically, with exception of phototrophic bacteria such as *Rhodocista pekingensis* 3-p^T (Zhang *et al.* 2003), which displays anaerobic growth.

Strain R5913^T grew between 10-36 °C (optimum at 27-33 °C), while strain R5959^T preferred a slightly higher temperature in the range of 24-40 °C (optimum at 33-36 °C). The results are in accordance with the temperature ranges of some closest relatives that grew optimally at approximately 30 °C (Table 10), including *Oceanibaculum pacificum* MCCC 1A02656^T (Dong *et al.* 2010) and *Dongia mobilis* CGMCC 1.7660^T (Liu *et al.* 2010). Strain R5913^T grew between pH 6.0 to 8.4 (optimum at pH 6.5-7.3) while strain R5959^T preferred a slightly lower pH in the range of 5.1-8.2 (optimum at pH 5.6-7.4). The majority of the closest genera have a higher alkalinity tolerance (Table 10), especially *Dongia rigui* 04SU4-P^T (Baik *et al.* 2013) and *Oceanibaculum indicum* P24^T (Lai *et al.* 2009) which can tolerate pH values of up to 11. Both strains grew optimally without the addition of NaCl. Strain R5959^T could tolerate a concentration of NaCl up to 0.5% while strain R5913^T grew slowly in 1% NaCl but could not grow in 3% NaCl. This trait is shared within some of their closest genera that were isolated mostly from soil and freshwater lakes, including *Reyranella* (Pagnier *et al.* 2011; Kim *et al*. 2013; Lee & Whang 2014; Lee *et al.* 2017; Cui *et al.* 2017), *Dongia* (Kim *et al.* 2016; Liu *et al.* 2010; Baik *et al.* 2013), *Aliidongia* (Chen *et al.* 2017), *Stella* (Vasilyeva, 1985), *Lacibacterium* (Sheu *et al.* 2013) and *Elstera* (Rahalkar *et al.* 2012; Cai *et al.* 2017). Nevertheless, some members of *Oceanibaculum* and *Nisaea*, which were isolated from marine environments, can tolerate up to 6 and 9% (w/v) NaCl, respectively (Dong *et al.* 2010; Lai *et al.* 2009; Du *et al.* 2017; Urios *et al.* 2008). Moreover, the doubling time of strains R5913^T and R5959^T under optimal conditions was 25.1 h and 31.7 h, respectively. The doubling time of most of the closest relatives has not been reported, except for *Elstera litoralis* Dia-1^T which grew slower (40 h) in VM-ethanol medium when compared to both strains (Rahalkar *et al.* 2012).

Both strains were tested positive for esterase, leucine arylamidase, acid phosphatase, and napthol-AS-BI-phosphohydrolase, and weakly positive for alkaline phosphatase, which are similar to those for *Tagaea marina* TT1^T, *Aliidongia dinghuensis* 7M-Z19^T, *Lacibacterium aquatile* LTC-2^T, *Elstera litoralis* Dia-1^T, and *Elstera cyanobacterium* TH019^T (Table 11). The majority of the closest relatives were tested positive for alkaline phosphatase but this was only weakly detected in both strains. Strain R5959^T tested positive for esterase lipase and valine arylamidase but these were only weakly detected in strain R5913^T.

Table 11 – Differential enzyme activities of strains R5913^T and R5959^T in comparison to other related taxa in the *Rhodospirillaceae* family.

Strains/genera: 1, R5913^T; 2, R5959^T; 3, *Reyranella* (Pagnier *et al.* 2011; Kim *et al*. 2013; Lee & Whang 2014; Lee *et al.* 2017; Cui *et al.* 2017); 4, *Tagea* (Jean *et al.* 2016); 5, *Oceanibaculum* (Dong *et al.* 2010; Lai *et al.* 2009; Du *et al.* 2017); 6, *Nisaea* (Urios *et al.* 2008); 7, *Thalassobaculum* (Zhang *et al.* 2008; Urios *et al.* 2010; Su *et al.* 2016); 8, *Dongia* (Kim *et al.* 2016; Liu *et al.* 2010; Baik *et al.* 2013); 9, *Inquilinus* (Jung *et al.* 2011; Coenye *et al.* 2002); 10, *Aliidongia* (Chen *et al.* 2007); 11, *Stella* (Vasilyeva 1985); 12, *Lacibacterium* (Sheu *et al.* 2013) ; 13, *Elstera* (Rahalkar *et al.* 2012; Cai *et al.* 2017); +, positive; -, negative; V, variable response; NA, no data available.

Enzyme activities	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
Alkaline phosphatase	w	w	+	+	+	+	V	V	NA	+	NA	+	+
Esterase (C4)	+	+	+	+	+	-	V	+	NA	+	NA	+	+
Esterase Lipase (C8)	w	+	+	+	V	-	V	+	NA	+	NA	+	+
Lipase (C14)	-	-	-	-	V	-	V	V	+	-	NA	+	-
Leucine arylamidase	+	+	V	+	V	+	+	+	NA	+	NA	+	+
Valine arylamidase	w	+	V	+	+	-	+	V	NA	+	NA	+	-
Cystine arylamidase	-	w	V	+	+	-	V	+	NA	-	NA	+	-
Trypsin	-	-	V	+	V	-	-	V	NA	-	NA	+	-
α-chymotrypsin	-	-	V	w	V	-	V	-	NA	-	NA	-	V
Acid phosphatase	+	+	+	w	+	+	+	V	NA	+	NA	+	+
Naphthol-AS-BI-phosphohydrolase	+	+	V	w	V	-	V	V	NA	+	NA	+	+
α-galactosidase	-	-	-	-	-	-	V	-	NA	-	NA	-	V
β-galactosidase	-	-	V	-	-	-	V	-	+	-	NA	-	V
β-glucuronidase	-	-	-	-	-	-	-	-	NA	-	NA	-	-
α-glucosidase	-	-	-	+	-	-	V	-	-	-	NA	-	-
β-glucosidase	-	-	V	w	-	-	V	V	V	-	NA	-	-
N-acetyl-β-glucosaminidase	-	-	-	-	-	-	V	-	+	-	NA	-	V
α-mannosidase	-	-	-	-	-	-	V	-	NA	-	NA	-	-
α-fucosidase	-	-	-	-	-	-	V	-	NA	-	NA	-	-
Indole production	-	-	-	NA	-	NA	-	-	-	-	NA	-	-
Arginine dihydrolase	-	-	V	NA	V	NA	-	-	-	NA	NA	-	NA
Urease	-	-	V	+	V	NA	+	V	-	+	NA	-	NA
Esculin hydrolysis (β-glucosidase)	-	-	V	+	-	NA	-	V	NA	-	NA	-	NA
Gelatin hydrolysis (protease)	-	-	V	+	-	NA	+	V	-	-	-	-	V
β-galactosidase (Para-NitroPhenyl-βD-Galactopyranosidase)	-	-	V	NA	-	NA	-	NA	NA	NA	NA	-	NA

Cystine arylamidase was weakly detected only in R5959^T. No activities were detected for lipase, trypsin, α-chymotrypsin, α-galactosidase, β-galactosidase, β-glucuronidase, α-glucosidase, β-glucosidase, N-acetyl-β-glucosaminidase, α-mannosidase, and α-fucosidase for both strains, similar to what have been observed for *Aliidongia dinghuensis* 7M-Z19^T (Table 11). In addition, strains R5913^T and R5959^T were tested negative for nitrate reduction and glucose fermentation. The capability to reduce nitrate varies in the *Rhodospirillaceae* family and only members of *Lacibacterium* and *Elstera* are able to ferment glucose (Table 10). Moreover, indole production, arginine dihydrolase, urease and β-galactosidase activities, as well as hydrolysis of esculin and gelatin were negative for both strains. The closest relatives of both strains were not able to produce indole (Table 11) while arginine dihydrolase were mostly negative and only present in *Reyranella* and *Oceanibaculum*.

Strain R5913^T metabolized yeast extract and grew weakly on casein hydrolysate, pepton, β-hydroxybutyrate, trimethoxybenzoate, acetate, ethylene glycol, butyrate, glycerol, α-hydroxybutyrate, sodium pyruvate, and isovaleric. Strain R5959^T only utilized yeast extract and grew weakly on casein hydrolysate. Similarly, a narrow range of growth substrates has been observed for *Tagaea marina* TT1^T (Jean *et al.* 2016) and *Thalassobaculum litoreum* DSM 18839^T (Zhang *et al.* 2008). The first tests negative for all substrates in Biolog GN2 assay and the latter is only able to utilize D-ribose, L-arabinose, sucrose, and yeast extract. Nevertheless, most of the closest relatives of R5913^T and R5959^T utilize a wider range of substrates (Table 12). The abilities to utilize complex substrates such as starch, cellulose, carboxymethyl (CM) cellulose, chitin, lignin, polygalacturonic acid, pectin, xylan and 2,2'-azinobis (3-ethylbenzthiazoline-6-sulfonate) (ABTS) were further tested in solidified R2A medium. Both strains tested negative for the utilization of all substrates while some members of *Reyranella* are able to utilize cellulose (Table 12).

The predominant fatty acids of strains R5913^T and R5959^T were $C_{19:0}$ cyclo *ω8c* (43.0 and 53.1%, respectively) and $C_{16:0}$ (15.0 and 25.1%, respectively) (Table 13). Contrastingly, $C_{18:1}$ *ω7c* (27.2%) was found as one of the major fatty acids of strain R5913^T but was much less abundant in strain R5959^T (1.6%). These three fatty acids are commonly reported as the major fatty acids of the closely related genera, including *Reyranella*, *Oceanibaculum*, *Thalassobaculum* and *Dongia* (Table 10) but $C_{19:0}$ cyclo *ω8c* and $C_{18:1}$ *ω7c* are absent in several genera, such as *Constrictibacter* and *Magnetospirillum* (Yamada *et al.* 2011; Dziuba *et al.* 2016). In addition, both strains also contained $C_{18:0}$ (1.9 and 5.2%, respectively) and $C_{16:0}$ 2-OH (1.7 and 5.1%, respectively). Fatty acids $C_{18:1}$ 2-OH (5.1%) and $C_{14:0}$ 2-OH (4.6%) were detected in R5913^T while traces were found in R5959^T (0.5 and 0.7%, respectively) and $C_{18:0}$ 3-OH (5.2%) was only detected in R5959^T (Table 10).

Table 12 – Utilization of carbon sources of strains R5913^T and R5959^T.

Strain: 1, R5913^T; 2, R5959^T; +, positive; -, negative; (+), weakly positive.

Substrates	1	2	Substrates	1	2	Substrates	1	2	Substrates	1	2
Arabinose	-	-	Xylitol	-	-	β-Hydroxybutyrate	(+)	-	2-Oxovalerate	-	-
Cellobiose	-	-	Alanine	-	-	γ-Hydroxybutyrate	-	-	Butanol	-	-
Erythrose	-	-	Arginine	-	-	Isobutyrate	-	-	1,2-Butandiol	-	-
Erythrulose	-	-	Asparagine	-	-	Caproate	-	-	2,3-Butandiol	-	-
Fructose	-	-	Aspartate	-	-	Caprylate	-	-	Ethanol	-	-
Fucose	-	-	Cysteine	-	-	Citrate	-	-	Ethylene glycol	(+)	-
Galactose	-	-	Glutamate	-	-	Isocitrate	-	-	Glycerol	(+)	-
Glucose	-	-	Glutamine	-	-	Crotonate	-	-	Methanol	-	-
Lactose	-	-	Glycine	-	-	Formate	-	-	Propanol	-	-
Lyxose	-	-	Histidine	-	-	Fumarate	-	-	1,2-Propandiol	-	-
Maltose	-	-	Hydroxy-Proline	-	-	Gluconate	-	-	Tween 80	-	-
Mannose	-	-	Isoleucine	-	-	2-Oxogluconate	-	-	Casamino acids	-	-
Melezitose	-	-	Leucine	-	-	Glucuronate	-	-	Casein hydrolysate	(+)	(+)
Raffinose	-	-	Lysine	-	-	2-Oxoglutarate	-	-	Peptone	(+)	-
Rhamnose	-	-	Methionine	-	-	Glycolate	-	-	Yeast extract	+	+
Sorbose	-	-	Ornithine	-	-	Glyoxylate	-	-	Laminarin	-	-
Sucrose	-	-	Phenylalanine	-	-	Heptanoate	-	-	2,2’-azinobis (3-ethylbenzthiazoline-6-sulfonate) (ABTS)	-	-
Trehalose	-	-	Proline	-	-	Isovaleric acid	(+)	-			
Xylose	-	-	Serine	-	-	Laevulinate	-	-			
Glucosamine	-	-	Threonine	-	-	Lactate	-	-	Carboxymethyl cellulose (CMC)	-	-
N-acetylglucosamine	-	-	Tryptophan	-	-	Malate	-	-			
N-acetylgalactosamine	-	-	Tyrosine	-	-	Maleic acid	-	-	Cellulose	-	-
Acetoin	-	-	Valine	-	-	Malonate	-	-	Chitin	-	-
Adonitol	-	-	Adipate	-	-	Nicotinic acid	-	-	Lignin	-	-
Arabitol	-	-	Acetate	(+)	-	Oxaloacetate	-	-	Polygalacturonic acid	-	-
Dulcitol	-	-	Ascorbate	-	-	Propionate	-	-	Pectin	-	-
Lyxitol	-	-	Benzoate	-	-	Pyruvate	(+)	-	Starch	-	-
Mannitol	-	-	Trimethoxybenzoate	(+)	-	Shikimate	-	-	Xylan	-	-
Myo-Inositol	-	-	Butyrate	(+)	-	Succinate	-	-			
Sorbitol	-	-	α-Hydroxybutyrate	(+)	-	Tartrate	-	-			

Table 13 – Fatty acid profiles of the novel strains. Summed features include at least two fatty acids that could not be distinguished by MIDI chromatograph. Numbers in bold reflect major fatty acids.

Fatty acid	Percentage of fatty acids	
	R5913^T	R5959^T
Saturated		
15:0	0.14	-
16:0	**15.03**	**25.14**
17:0	0.09	0.31
18:0	1.86	**5.20**
Unsaturated		
18:1 ω7c	**27.24**	1.56
18:1 ω9c	-	0.17
20:2 ω6,9c	-	0.45
11 methyl 18:1 ω7c	0.39	0.59
Cyclopropane acids		
19:0 cyclo ω8c	**42.95**	**53.16**
Hydroxy		
12:0 2-OH	0.12	-
14:0 2-OH	4.63	0.74
16:0 2-OH	1.65	**5.09**
16:0 3-OH	0.35	0.15
17:0 3-OH	-	0.11
18:0 2-OH	0.17	0.42
18:0 3-OH	-	**5.15**
18:1 2-OH	**5.07**	0.53
Summed Feature		
3 (16:1 ω7c/15 iso 2OH)	0.13	-
4 (17:1 iso I/anteiso B)	0.09	0.12
Unknown		
ECL* 14.959	-	1.10
ECL* 18.814	0.11	-

*ECL, equivalent chain-length.

The predominant polar lipids of strain R5913^T were phosphatidylethanolamine, an unknown aminoglycolipid and two unidentified aminoglycophospholipids (Figure 57). For strain R5959^T, these included phosphatidylethanolamine, an unidentified aminolipid, an unknown phosphoglycolipid, an unknown aminoglycophospholipids, and three unidentified aminoglycolipids (Figure 58). Phosphatidylethanolamine and phosphatidylglycerol are commonly reported for members of the *Rhodospirillaceae* family (Table 10), but the latter was not identified in both novel strains. Ubiquinone-10 was the sole respiratory quinone for both strains, and it is also the major quinone found in their phylogenetically closest genera (Table 10).

The maximum likelihood and Neighbor Joining trees confirmed the affiliation of the strain R5913^T and R5959^T within the family *Rhodospirillaceae* (Figure 59 and Figure 60, respectively). Both strains formed a well-defined monophyletic group, that is supported by a high bootstrap value. In both phylogenetic trees, *Reyranella* was the phylogenetically closest genus of the new isolates with a high statistical support.

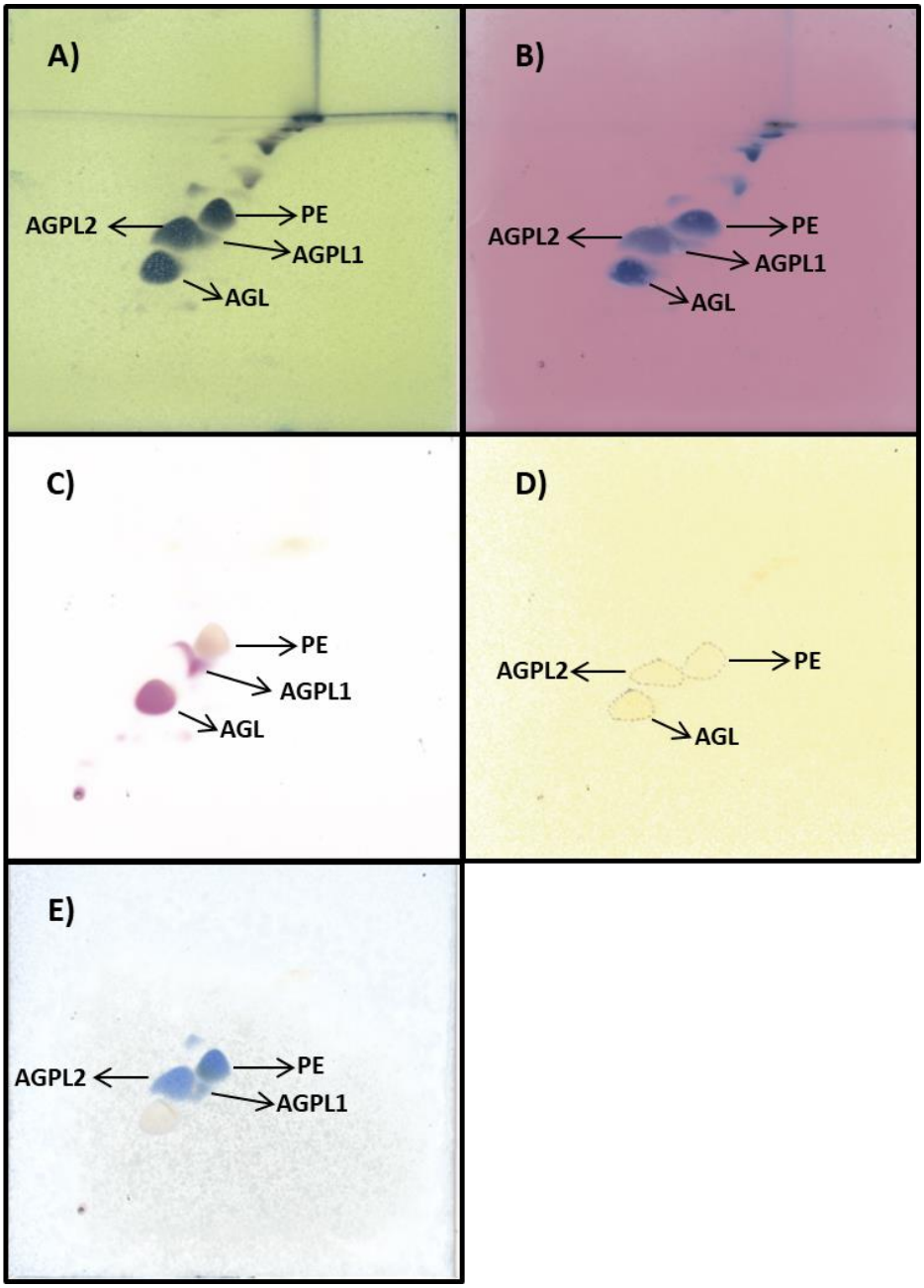


Figure 57 – Polar lipid profiles of strain R5913^T based on thin layer chromatography. Staining for determination of polar lipids with dodecamolybdophosporic acid (A), anisaldehyde sulfuric acid (B), ninhydrin (C), dragendorff (D), and molybdenum blue (E). Lipid separation was achieved by chloroform:methanol:water (65:25:4, v/v/v) in the first direction and chloroform:methanol:acetic-acid:water (80:12:15:4, v/v/v/v) in the second direction. PE: phosphatidylethanolamine, AGL: an unidentified aminoglycolipid, AGPL: an unidentified aminoglycophospholipid.

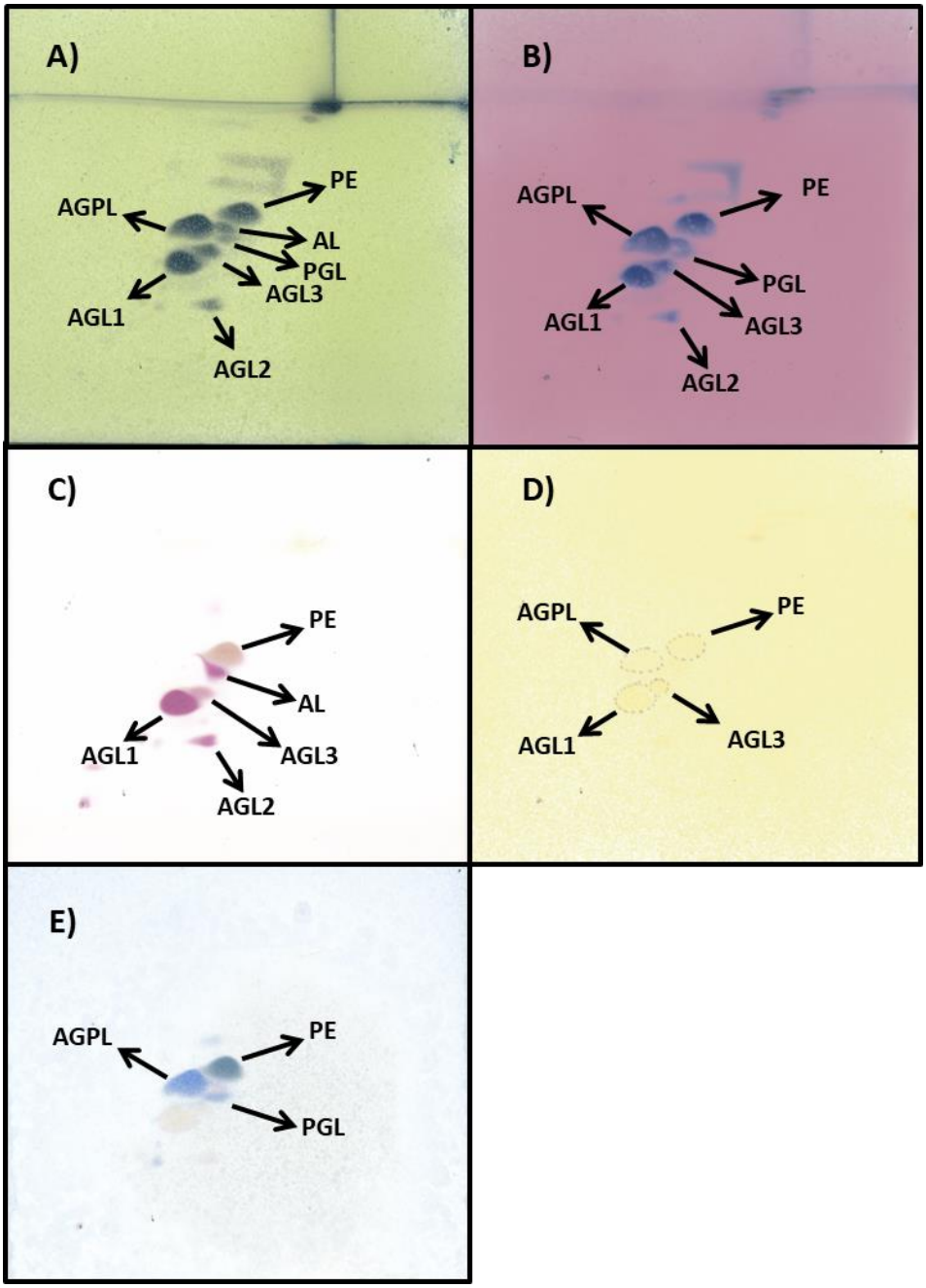


Figure 58 – Polar lipid profiles of strain R5959^T on thin layer chromatography. Staining for determination of polar lipids with dodecamolybdophosporic acid (A), anisaldehyde sulfuric acid (B), ninhydrin (C), dragendorff (D), and molybdenum blue (E). Lipid separation was achieved by chloroform:methanol:water (65:25:4, v/v/v) in the first direction and chloroform:methanol:acetic-acid:water (80:12:15:4, v/v/v/v) in the second direction. PE: phosphatidylethanolamine, AGL: an unidentified aminoglycolipids, PGL: an unidentified phosphoglycolipid, AGPL: an unidentified aminoglycophospholipid, AL: aminolipid.

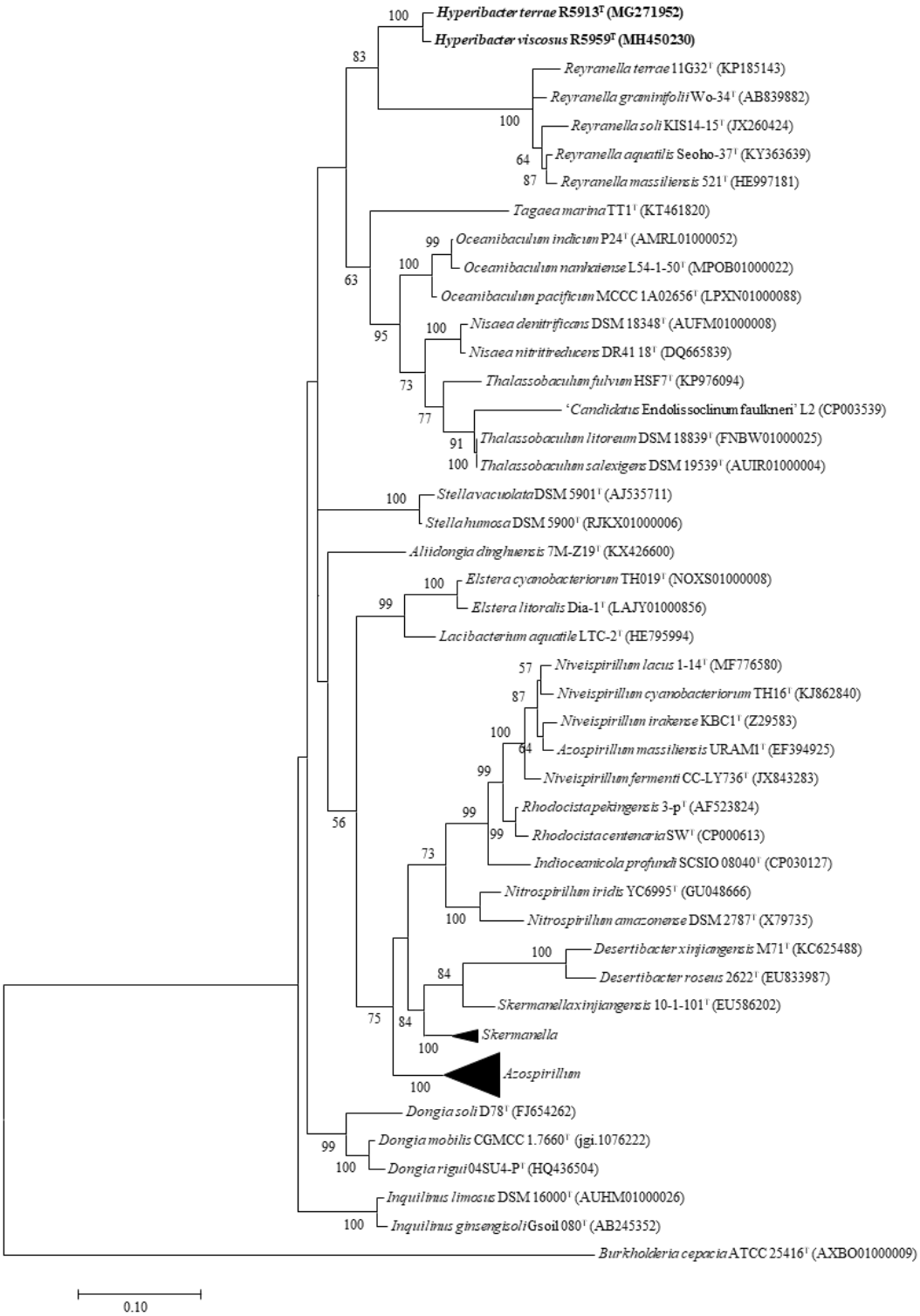


Figure 59 – Maximum-likelihood (ML) phylogenetic tree based on the almost full-length 16s rRNA gene sequences illustrating the position of strains R5913^T and R5959^T in the *Rhodospirillaceae* family. The best evolutionary model for nucleotide substitution, calculated by Mega 7.0, and applied for the tree was T92+G+I. Scale bar, 0.1 fixed nucleotide substitutions per site. *Burkholderia cepacia* ATCC 25416^T was used as an outgroup. Only bootstrap values above 50% (of 1000 replicates) are indicated at the tree branching points.

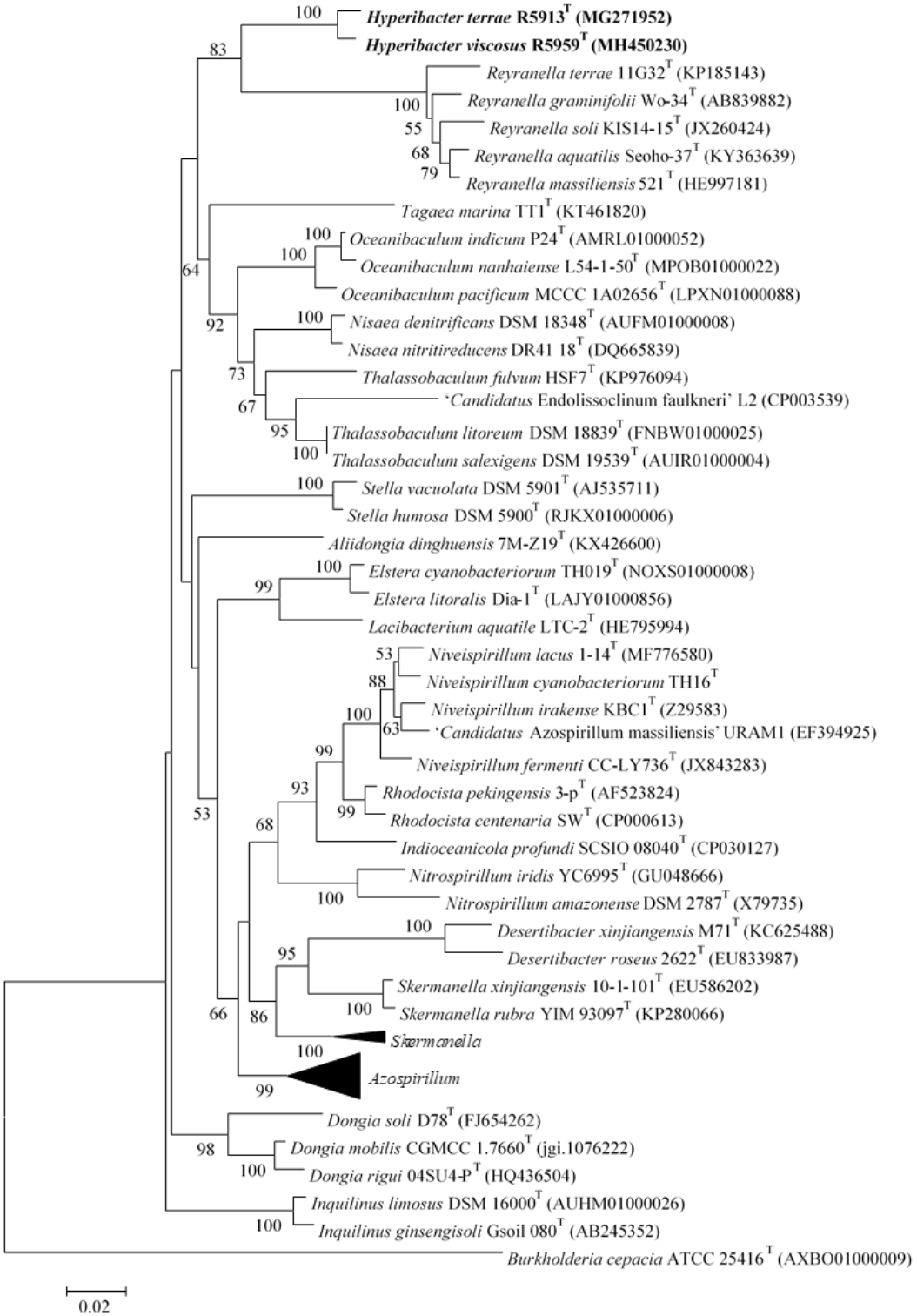


Figure 60 – Neighbor-joining (NJ) phylogenetic tree based on the almost full-length 16S rRNA gene sequences illustrating the position of strains R5913^T and R5959^T in the *Rhodospirillaceae* family. Scale bar, 0.02 nucleotide divergence. Kimura 2-parameter model was employed to calculate the evolutionary distances. *Burkholderia cepacia* ATCC 25416^T was chosen as an outgroup. Only bootstrap values above 50% (of 1000 replicates) indicated at the branching points.

5.1.4.2 – Genomic features of strains R5913^T and R5959^T

The genomes of strains R5913^T and R5959^T were organized in one circular chromosome with the size and G+C content of 5,894,118 bp and 66.0%, and 5,865,246 bp and 67.4%, respectively. The number of predicted genes of strains R5913^T and R5959^T were 5459 and 5336, comprising 5361 and 5247 protein-coding sequences, and 56 and 53 tRNAs, respectively. Both strains had 6 rRNA genes organized in two operons (Figure 61).

The average nucleotide identity (ANI) value between the two novel strains, calculated with OrthoANIu algorithm using EzGenome web service (www.ezbiocloud.net/tools/ani) (Yoon *et al.* 2017b), was 83.8%. The ANI values calculated between the strains and their closest genera were lower. Strains R5913^T and R5959^T shared ANI values of 72.0% and 72.2% with *Dongia mobilis* CGMCC 1.7660^T and 70.8% and 71.0% with *Oceanibaculum pacificum* MCCC 1A02656^T, respectively. Confirming the results of the 16S rRNA-based analysis, both strains form a monophyletic group in the phylogenomic tree, with *Dongia mobilis* CGMCC 1.7660^T being the closest neighbor (Figure 62). In addition, the branching pattern was supported by high bootstrap values and gene support indices.

On the basis of phenotypic, genomic, and phylogenetic investigations, the affiliation of both strains within the family *Rhodospirillaceae* is confirmed. Both strains could be differentiated by cell morphology, enzymatic activities, optimal growth conditions including temperature and pH, substrate preferences, polar lipids and fatty acids compositions, and thus should be assigned to different species, as also supported by low ANI value (83.8%) which is below the threshold for the delineation of bacterial species (95.0 - 96.0%) (Richter & Roselló-Móra 2009).

5.1.4.3 – Description of *Hypericibacter* gen. nov.

Hypericibacter (Hy.pe.ri.ci.bac'ter. L. neut. n. *Hypericum* a botanical genus; N. L. masc. n. *bacter* a rod; a rod-shaped bacterium isolated from the rhizosphere of *Hypericum perforatum*).

Gram negative, motile by means of a single polar flagellum, non-spore-forming, non-capsulated, straight to slightly curved, short, rod-shaped bacteria that divide by binary fission. Oxidase positive, grow under aerobic and microaerophilic conditions, mesophilic, chemo-organotroph, which are unable to reduce nitrate or ferment glucose. The sole respiratory quinone is Q-10 and the predominant fatty acids are $C_{19:0}$ cyclo ω8c and $C_{16:0}$. The type strain is *Hypericibacter terrae.*

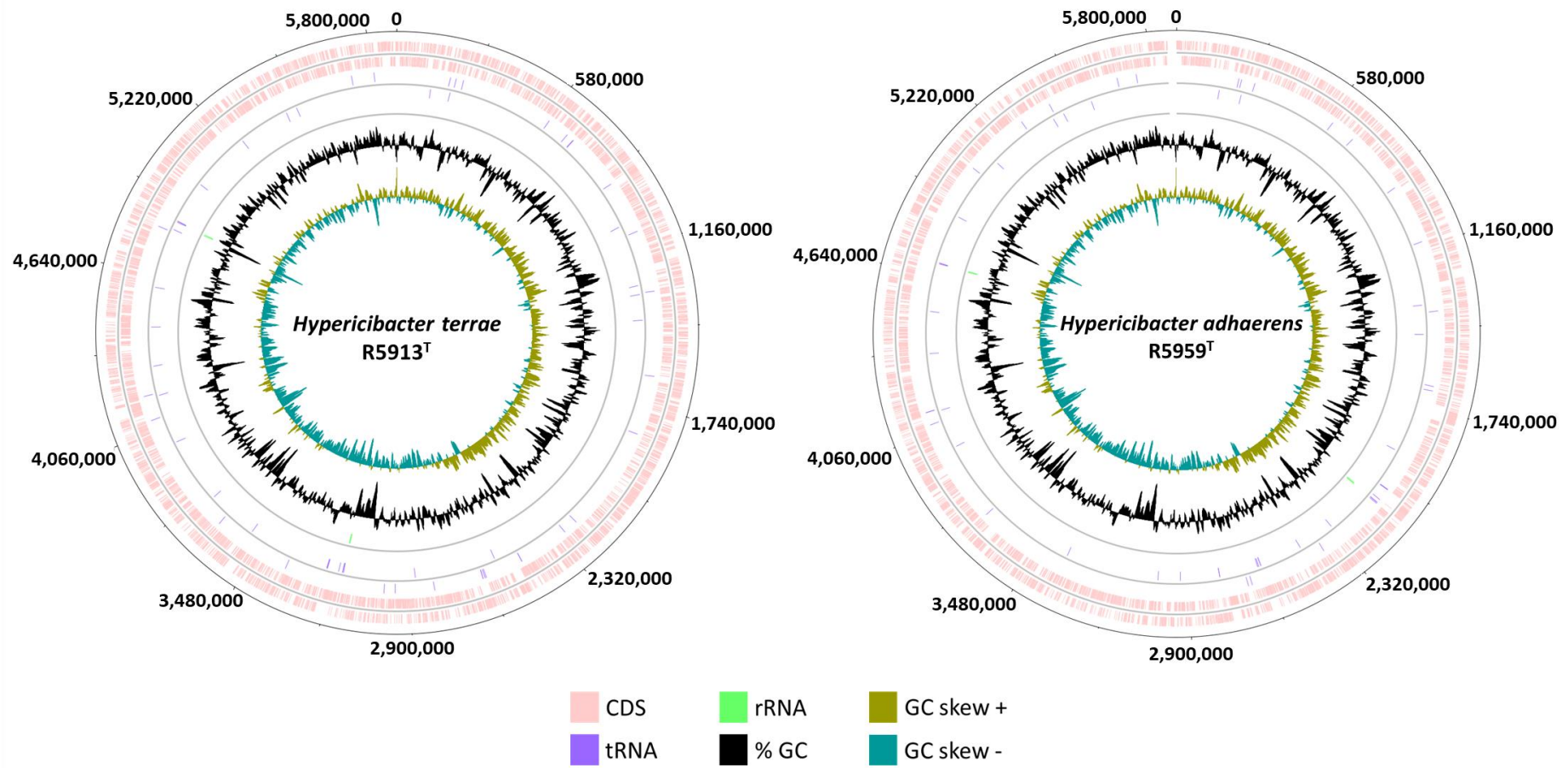


Figure 61 – Genome profiles of strains R5913^T (left) and R5959^T (right), visualized by DNA plotter in Artemis (Carver *et al.* 2012). Grey circle separates forward from reverse reads.

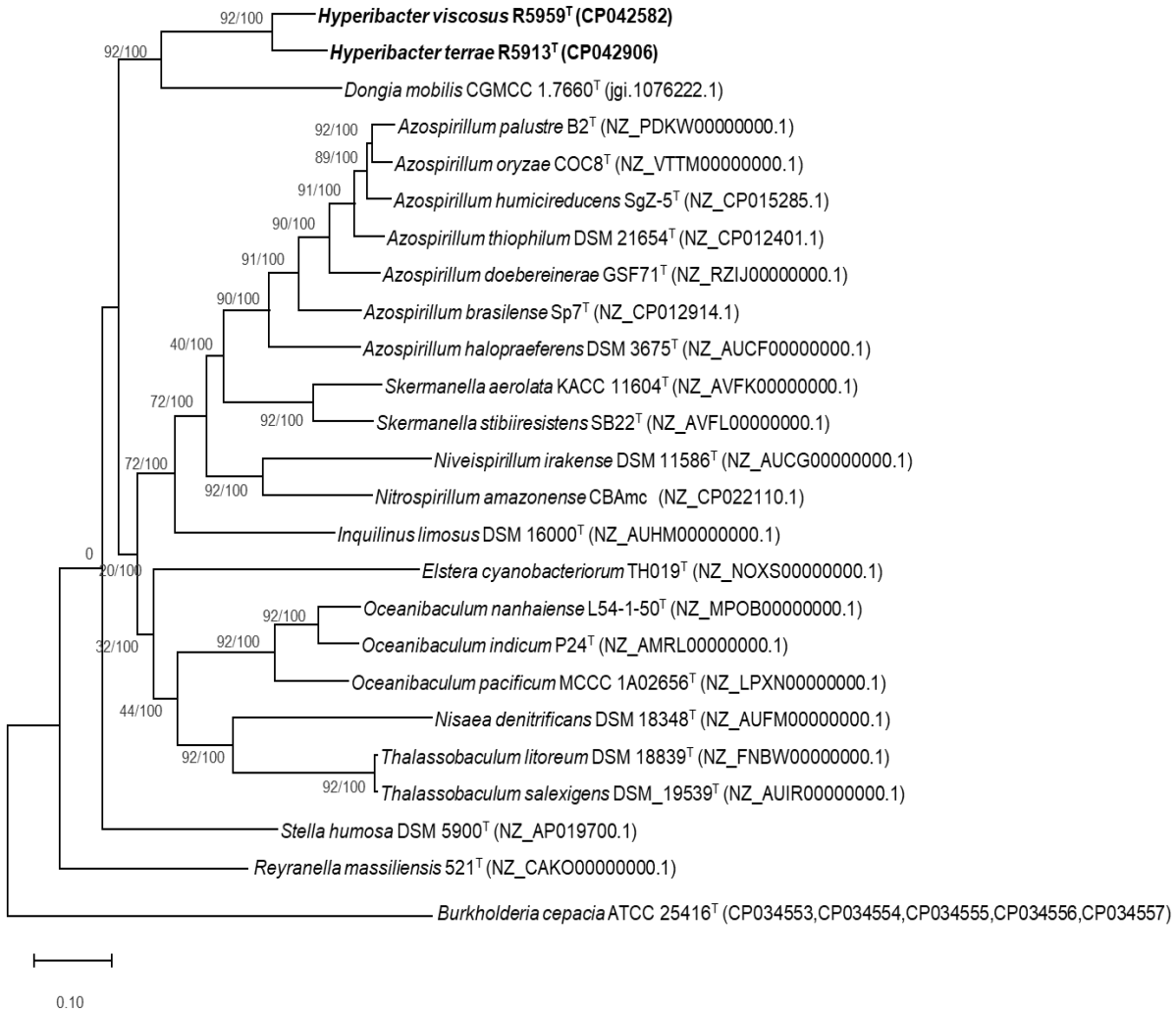


Figure 62 – Phylogenomic tree of strains R5913^T and R5959^T with their closest relatives in the family *Rhodospirillaceae*. The unrooted maximum likelihood phylogenetic tree based on a multiple alignment of a set of 92 gene (nucleotidic) sequences by the UBCG v. 3.0 pipeline. Bootstrap analysis was performed with 100 replications. Gene support indices (max. value; 92 genes) and percentage bootstrap values (max. value; 100%) are given at the tree branching points. Scale bar, 0.10 substitution per position.

5.1.4.4 – Description of *Hypericibacter terrae* sp. nov.

Hypericibacter terrae (ter´rae. L. gen. n. *terrae* of the soil, referring to the isolation source of the type strain.)

Exhibits the following characteristics in addition to those given in the genus description. Cells are 0.7-1.5 µm long and 0.5-0.8 µm wide and appear mostly as single cells. Colonies in solidified R2A medium are small, round, smooth, convex, and shiny white with a diameter of approximately 1 mm. Growth occurs between 20 and 36 °C (optimum 27-33 °C), between pH 6.0 and 8.4 (6.5-7.5), and at

up to 1% (w/v) NaCl (optimally without NaCl). The doubling time under optimal conditions is 25.1 h. Assimilates yeast extract but grows weakly on casein hydrolysate, pepton, β-hydroxybutyrate, trimethoxybenzoate, acetate, ethylene glycol, butyrate, glycerol, α-hydroxybutyrate, Na pyruvate, and isovaleric acid. No growth is observed on glucose, lactose, fructose, cellobiose, galactose, mannose, melezitose, raffinose, fucose, sorbose, lyxose, maltose, rhamnose, sucrose, trehalose, xylose, adonitol, arabitol, mannitol, *myo*-inositol, sorbitol, xylitol, lysin, hydroxyproline, casamino acid, glycolate, malonate, propionate, oxaloacetate, lactate, butanol, ethanol, methanol, propanol, *N*-acetylglucosamine, caproate, caprylate, dulcitol, erythrose, erythrulose, isocitrate, laevulinate, arabinose, glucosamine, gluconate, glucuronate, lyxitol, 2-oxoglutarate, acetoin, ascorbate, glyoxylate, 2-oxovalerate, 2-oxogluconate, *N*-acetylgalactosamine, maleic acid, 1,2-butandiol, 2,3-butandiol, 1,2-propandiol, alanine, arginine, asparagine, cysteine, glutamine, isoleucine, ornithine, proline, benzoate, tryptophan, formate, γ-hydroxybutyrate, isobutyrate, tyrosine, serine, phenylalanine, glycine, leucine, histidine, valine, methionine, threonine, succinate, nicotinic acid, tween 80, adipate, shikimate, aspartate, glutamate, laminarin, malate, citrate, tartrate, heptanoic acid, and fumarate. Tests positive for esterase, leucine arylamidase, acid phosphatase, and Naphthol-AS-BI-phosphohydrolase. Alkaline phosphatase, esterase lipase, and valine arylamidase acitivities are detected weakly. No activities of lipase, cystine arylamidase, trypsin, α-chymotrypsin, α-galactosidase, β-galactosidase, β-glucuronidase, α-glucosidase, β-glucosidase, *N*-acetyl-β-glucosaminidase, α-mannosidase, and α-fucosidase can be detected. Tests negative for nitrate reduction, indole production, and glucose fermentation. Activities of urease, arginine dihyrolase, and β-galactosidase (PNPG) are not present and aesculin and gelatin are not hydrolyzed.

The type strain is R5913^T (=DSM 109816^T = CECT 9472^T), isolated from the rhizosphere of *Hypericum perforatum* taken from the greenhouse of the Leibniz Institute of Plant Biochemistry (IPB), Halle, Germany (51° 29' 42.23" N, 11° 56' 36.56" E). The genomic G+C content of the type strain is 66.0%. The almost full length 16S rRNA gene and whole-genome sequences have been deposited in GenBank/EMBL/DDBJ under accession numbers MG271952 and CP042906, respectively.

5.1.4.5 – Description of *Hypericibacter adhaerens* sp. nov.

Hypericibacter adhaerens (ad.haeˊrens. L. part. adj. *adhaerens* sticky, referring to the sticky attribute of the colonies to solid growth media).

Exhibits the following characteristics in addition to those given in the genus description. Cells are 0.5-2.5 µm long and 0.3-1.0 µm wide and sometimes appear in short chains. Colonies in solidified R2A are small, round, smooth, convex, and shiny white with a diameter of approximately 1 mm.

Growth occurs between 24 and 40 °C (optimum 33 to 36 °C), between pH 5.0 and 8.2 (optimum 5.6 to 7.4), and up to 0.5% (w/v) NaCl (optimally without NaCl). The doubling time under optimal condition is 31.7 h. Grows on yeast extract but weakly on casein hydrolysate. No growth detected on glucose, lactose, fructose, cellobiose, galactose, mannose, melezitose, raffinose, fucose, sorbose, lyxose, maltose, rhamnose, sucrose, trehalose, xylose, adonitol, arabitol, mannitol, *myo*-inositol, sorbitol, xylitol, lysin, hydroxy-proline, casamino acid, glycolate, malonate, propionate, oxaloacetate, lactate, butanol, ethanol, glycerol, methanol, propanol, pepton, *N*-acetylglucosamine, caproate, caprylate, dulcitol, ethylenglycol, erythrose, erythrulose, α-hydroxybutyrate, isocitrate, laevulinate, arabinose, glucosamine, gluconate, glucuronate, lyxitol, 2-oxoglutarate, Na pyruvate, acetoin, ascorbate, glyoxylate, 2-oxovalerate, 2-oxogluconate, *N*-acetylgalactosamine, maleic acid, 1,2-butandiol, 2,3-butandiol, 1,2-propandiol, alanine, arginine, asparagine, cysteine, glutamine, isoleucine, ornithine, proline, benzoate, tryptophane, acetate, butyrate, formate, β-hydroxybutyrate, γ-hydroxybutyrate, isobutyrate, tyrosine, serine, phenylalanine, glycine, leucine, histidine, valine, methionine, threonine, succinate, nicotinic acid, tween 80, adipate, shikimate, aspartate, glutamate, laminarine, malate, citrate, tartrate, isovaleric acid, heptanoic acid, fumarate, and trimethoxybenzoate. Tests positive for esterase, esterase lipase, leucine arylamidase, valine arylamidase, acid phosphatase, and Napthol-AS-BI-phosphohydrolase, while alkaline phosphatase and cystine arylamidase activities are detected but weakly. No activities of lipase, trypsin, α-chymotrypsin, α-galactosidase, β-galactosidase, β-glucuronidase, α-glucosidase, β-glucosidase, *N*-acetyl-β-glucosaminidase, α-mannosidase, and α-fucosidase can be detected. Nitrate reduction, indole production, and glucose fermentation are all negative. Both aesculin and gelatin are not hydrolysed. Urease, arginine dihydrolase, and β-galactosidase (PNPG) activities are not present.

The type strain is R5959^T (=DSM 109817^T = CECT 9620^T), isolated from the rhizosphere of *Hypericum perforatum* taken from the greenhouse of the Leibniz Institute of Plant Biochemistry (IPB), Halle, Germany (51° 29' 42.23" N, 11° 56' 36.56" E). The genomic G+C content of the type strain is 67.4%. The almost full-length 16S rRNA gene and whole-genome sequences have been deposited in GenBank/EMBL/DDBJ under accession numbers MH450230 and CP042582, respectively.

5.2 – Discussion

5.2.1 – High throughput cultivation with low nutrient media and dilution to extinction principle to retrieve novel bacterial taxa

In 2016, a cultivation effort was conducted to assist a study on bacterial communities associated to *Hypericum* plants. High throughput was performed to allow isolation of higher number of soil taxa (e.g. from the rhizosphere of *H. perforatum*) that are known to be very diverse, where up to 53,000 bacterial species or 10^6 bacterial phylotypes can be found only in a single gram. However, the majority of soil bacteria are so far unculturable with only less than 1% can be cultured in the laboratory (Sandaa *et al.* 1996, Roesch *et al.* 2007; Overmann 2016). Thus, ecological functions of many soil taxa could not be fully determined as thorough investigation of soil bacteria requires isolates that can be cultured in the laboratory. Novel bacterial taxa with less than 97% similarity of 16S rRNA gene sequences with known isolates (and thus represent at least a novel species) were targeted in our cultivation attempt since they may perform novel/less studied functions related to the host plant, including production of hypericin and hyperforin with anti-depression effect in the rhizosphere of *Hypericum* plants.

Strategies that are employed include application of dilution to extinction principle and utilization of growth media with low-nutrient content (Connon & Giovannoni 2002), namely SSE/HD1:10 and R2A 1:10. The aim is to promote the growth of oligotrophic bacteria instead of fast-growing ones. The result confirms the effect of cultivation media on the diversity of grown taxa that has been previously observed (Davis *et al.* 2005). In addition, SSE/HD1:10 medium which is based on soil solution equivalent that mimics soil nutrient content (Huber *et al.* 2016) is more suitable for the growth of soil bacteria compared to R2A 1:10, as suggested with higher alpha diversity measures.

It is also interesting to note that some detected genera inside the wells of microtiter plates are not detected in the original soil samples, including *Aminobacter, Novosphingobium, Cutibacterium, Lysinimonas, and Nesterenkonia*. Instead of soil habitat, these genera are affiliated with mine tailings (Maynaud *et al.* 2012), groundwater (Tiirola *et al.* 2002), gut microbiota (Martin *et al.* 2018), tree bark (Kobayashi & Aoyagi 2019), and Antarctic desert soil (Aliyu *et al.* 2016). It is likely that the abundance of these genera in the original soil sample is very low and below the detection limit and hence escape detection. Since soil is a highly diverse habitat where living condition is affected by many biotic and abiotic factors, the low abundance of these genera may be caused by many factors, such as competition with other soil taxa, limited nutrient availability, unfavourable pH, and many others.

Some bacterial phyla with low numbers of representative isolates, such as Chloroflexi, Verrucomicrobiota, Planctomycetota, and Acidobacteriota are also detected in the liquid media. This implies that the employment of low-nutrient media has the potential to improve cultivation success, since it reduces cultivation bias toward Alphaproteobacteria and Gammaproteobacteria, Firmutes, Actinobacteria, and Bacteroidota (Overmann 2013). Unfortunately, many taxa could not be purified further mainly because they were overgrown by fast growing taxa when streaked onto a fresh agar medium. Completely avoiding the neighbor, by starting the cultivation with only 1 cell, is not an ideal solution as many unculturable bacteria rely on helper bacteria (Kaeberlein *et al.* 2002; Axelsson-Olsson *et al.* 2007). Small genome size is thought to be one of the determinants for the dependency, as more biological pathways will be lost as genome size decreases, supported by the fact that Candidate Phyla Radiation with small genome sizes still has no representative isolates so far (Hug *et al.* 2016). For this reason, co-cultivation should be employed in the coming cultivation efforts that aim toward retrieval of phylogenetically novel taxa.

5.2.2 – Co-cultivation approach is promising to isolate of novel taxa

The importance of co-cultivation is emphasized in the experiment with a novel member of Acidobacteriota which shares 96.7% similarity of 16S rRNA gene sequences with *Stenotrophobacter roseus* and thus represents a novel species in the genus *Stenotrophobacter*. The strain cannot be grown axenically and grows only when helper bacteria are present in close proximity. The effect of helper bacterium (*Pseudomonas prosekii*) supernatant on the growth of the Acidobacteriota strain is determined in liquid cultivation media. The result revealed that the strain growth is promoted with the supernatant, suggesting production of crucial compounds by the helper bacterium. The application of helper bacteria supernatant to promote growth of novel taxa has been previously reported, for example with the utilization of supernatant from a *Sphingomonas* strain culture to promote growth of *Catellibacter nectariphilum* and the utilization of *Geobacillus toebii* cell extract to isolate specific members of family *Clostridiaceae* (Tanaka *et al.* 2004; Kim *et al.* 2008; Kim *et al.* 2011; Stewart 2012).

The helper bacteria of the Acidobacteriota strain are not limited to a specific taxonomic rank, since members of *Pseudomonas, Bacillus,* and *Mesorhizobium* are able to promote the strain growth in solidified SSE/HD1:10. Those three taxa produce remarkably diverse secondary metabolites (Gross & Loper 2009; Sansinenea & Ortiz 2011; Peng *et al.* 2014), where some of the compounds are likely crucial for the growth of the Acidobacteriota strain. Interestingly, *Pseudomonas, Mesorhizobium*, and *Bacillus* are also observed to be positively correlated with 27, 26, and 12 other bacterial genera, respectively, based on the co-occurrence test where each single well of microtiter plates is

investigated using amplicon sequencing. This implies that diverse metabolites produced by the three genera might allow the growth of the neighboring taxa, independently of the phylogenetic affiliation.

Further investigation of the Acidobacteriota strain with the corresponding helper bacteria should be performed in the future to reveal the vital compounds. Metabolomic and genomic approaches can be performed simultaneously to identify the compounds and the missing biological pathways. The study will be advantageous to gain deeper understanding of unculturable taxa and to build new strategies that can be employed to improve cultivation success.

5.2.3 – Ecological functions of the novel isolates might be more prominent in other habitats

The two new members of *Rhodospirillaceae* can be considered as members of the rare biosphere (the abundance is very low, often below 0.01%) in their original habitat and thus act as a seed bank that can be active given favorable environmental conditions (Lennon & Jones 2011; Bickel & Or 2020). However, the abundance of these taxa can be higher than 1% in other habitats based on alignment to IMNGS depository. For example, strain R5913^T abundance can reach 2.4% in polluted-soil samples and 1.6% in the rhizosphere of willow trees, while strain R5959^T abundance can be up to 1.6% in the refused-dumps of leaf cutter ants and 0.6% in the oil-polluted soil. Further investigation of these taxa can be focused on these habitats in order to determine important ecological functions that these taxa might carry. Nonetheless, rare bacteria with low abundance still play important roles in maintaining community stability and in the turnover of specific elements, as depicted by members of *Desulfosporosinus* that reduce sulfate in a peatland (Pester *et al.* 2010).

This chapter aims to provide recommendations for the coming cultivation efforts targeting phylogenetically novel soil taxa, since the majority of soil bacteria still resist cultivation. The application of low-nutrient cultivation media and dilution to extinction principle with inoculation of less than 10 cells, are recommended to target oligotrophs instead of copiotrophs. In addition, employment of SSE/HD1:10 medium is observed to be better than R2A1:10 in recapturing soil diversity. Co-cultivation approach is also beneficial and should be utilized more often in the future. Moreover, further investigation of the Acidobacteriota strain with the helper bacteria might reveal crucial information that is needed to obtain the pure isolate of the target bacterium.

5.3 – References

Aliyu, H., De Maayer, P. & Cowan, D. 2016. The genome of the Antarctic polyextremophile *Nesterenkonia* sp. AN1 reveals adaptive strategies for survival under multiple stress conditions. *FEMS Microbiology Ecology*, **92**, doi: 10.1093/femsec/fiw032.

Axelsson-Olsson, D., Ellström, P., Waldenström, J., Haemig, P. D., Brudin, L. & Olsen, B. 2007. *Acanthamoeba-Campylobacter* Coculture as a Novel Method for Enrichment of *Campylobacter* Species. *Applied and Environmental Microbiology*, **73**, 6864-6869, doi: 10.1128/AEM.01305-07.

Baik, K. S., Hwang, Y. M., Choi, J. -S., Kwon, J. & Seong, C. N. 2013. *Dongia rigui* sp. nov., isolated from freshwater of a large wetland in Korea. *Antonie van Leeuwenhoek*, **104**: 1143-1150, doi: 10.1007/s10482-013-0036-9.

Bickel, S. & Or, D. 2020. Soil bacterial diversity mediated by microscale aqueous-phase processes across biomes. *Nature Communications*, **11**, doi: 10.1038/s41467-019-13966-w.

Cai, H., Zeng, Y., Wang, Y. & Jiang, H. 2017. *Elstera cyanobacteriorum* sp. nov., a novel bacterium isolated from cyanobacterial aggregates in a eutrophic lake. *International Journal of Systematic and Evolutionary Microbiology*, **67**, 4272-4275, doi: 10.1099/ijsem.0.002308.

Carver, T., Harris, S. R., Berriman, M., Parkhill, J. & McQuillan, J. A. 2012. **Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data.** *Bioinformatics, **28**, 464-469, doi:* 10.1093/bioinformatics/btr703.

Chen, M. H., Zhou, X. Y., Ou, F. H., Xia, F., Lv, Y. Y. & Qiu, L. H. 2017. *Aliidongia dinghuensis* gen. nov., sp. nov., a poly-ß-hydroxybutyrate-producing bacterium isolated from *Pinus massoniana* forest soil. *International Journal of Systematic and Evolutionary Microbiology*, **67**, 212-217, doi: 10.1099/ijsem.0.001588.

Coenye, T., Goris, J., Spilker, T., Vandamme, P. & LiPuma, J. J. 2002. Characterization of Unusual Bacteria Isolated from Respiratory Secretions of Cystic Fibrosis Patients and Description of *Inquilinus limosus* gen. nov., sp. nov. *Journal of Clinical Microbiology*, **40**, 2062-2069, doi: 10.1128/jcm.40.6.2062-2069.2002.

Connon, S. A. & Giovannoni, S. J. 2002. High-Throughput Methods for Culturing Microorganisms in Very-Low-Nutrient Media Yield Diverse New Marine Isolates. *Applied and Environmental Microbiology*, 2002, **68**, 3878-3885, doi: 10.1128/AEM.68.8.3878-3885.2002.

Cui, Y., Chun, S. -J., Ko, S. -R., Lee, H. G., Srivastava, A., Oh, H. -M. & Ahn, C. -Y. 2017. *Reyranella aquatilis* sp. nov., an alphaproteobacterium isolated from a eutrophic lake. *International Journal of Systematic and Evolutionary Microbiology*, **67**, 3496-3500, doi: 10.1099/ijsem.0.002151.

Davis, K. E. R., Joseph, S. J. & Janssen, P. H. 2005. Effects of Growth Medium, Inoculum Size, and Incubation Time on Culturability and Isolation of Soil Bacteria. *Applied and Environmental Microbiology*, **71**, 826-834, doi: 10.1128/AEM.71.2.826-834.2005.

Dong, C., Lai, Q., Chen, L., Sun, F., Shao, Z. & Yu, Z. 2010. *Oceanibaculum pacificum* sp. nov., isolated from hydrothermal field sediment of the south-west Pacific Ocean. *International Journal of Systematic and Evolutionary Microbiology*, **60**, 219-222, doi: 10.1099/ijs.0.011932-0.

Du, Y., Liu, X., Lai, Q., Li, W., Sun, F. & Shao, Z. 2017. *Oceanibaculum nanhaiense* sp. nov., isolated from surface seawater. *International Journal of Systematic and Evolutionary Microbiology*, **67**, 4842-4845, doi: 10.1099/ijsem.0.002388.

Dziuba, M., Koziaeva, V., Grouzdev, D., Burganskaya, E., Baslerov, R., Kolganova, T., Chernyadyev, A., Osipov, G., Andrianova, E., Gorlenko, V. & Kuznetsov, B. 2016. *Magnetospirillum caucaseum* sp. nov., *Magnetospirillum marisnigri* sp. nov. and *Magnetospirillum moscoviense* sp. nov., freshwater magnetotactic bacteria isolated from three distinct geographical locations in European Russia. *International Journal of Systematic and Evolutionary Microbiology*, **66**, doi: 10.1099/ijsem.0.000994.

Gross, H. & Loper, J. E. 2009. Genomics of secondary metabolite production by *Pseudomonas* spp. *Natural Product Reports*, **26**, 1408-1446, doi:10.1039/b817075b.

Huber, K.J., Geppert, A. M., Wanner, G., Fösel, B. U., Wüst, P. K. & Overmann, J. 2016. The first representative of the globally widespread subdivision 6 Acidobacteria, Vicinamibacter silvestris gen. nov., sp. nov., isolated from subtropical savannah soil. *International Journal of Systematic and Evolutionary Microbiology*, **66**, 2971–2979, doi: 10.1099/ijsem.0.001131.

Hug, L. A., Baker, **B. J., Anantharaman, K., Brown, C. T., Probst, A. J., Castelle, C. J., Butterfield, C. N., Hernsdorf, A. W., Amano, Y., Ise, K., Suzuki, Y., Dudek, N., Relman, D. A., Finstad, K. M., Amundson, R., Thomas, B. C. & Banfield**, J. F. 2016. A new view of the tree of life. *Nature Microbiology*, **1,** doi: 10.1038/nmicrobiol.2016.48

Jean, W. D., Huang, S. P., Chen, J. S. & Shieh, W. Y. 2016. *Tagea marina* gen. nov., sp. nov., a marine bacterium isolated from shallow coastal water. *International Journal of Systematic and Evolutionary Microbiology*, 66, 592-597, doi: 10.1099/ijsem.0.000756.

Jung, H. M., Lee, J. S., Bae, H. -M., Yi, T. -H., Kim, S. -Y., Lee, S. -T. & Im, W. -T. 2011. *Inquilinus ginsengisoli* sp. nov., isolated from soil of a ginseng field. *International Journal of Systematic and Evolutionary Microbiology*, **61**, 201-204, doi: 10.1099/ijs.0.018689-0.

Kaeberlein, T., Lewis, K. & Epstein, S. S. 2002. Isolating "uncultivable" microorganisms in pure culture in a simulated natural environment. *Science*, **296**, 1127-1129, doi: 10.1126/science.1070633.

Kim, D. -U., Lee, H., Kim, H., Kim, S. -G. & Ka, J. -O. 2016. *Dongia soli* sp. nov., isolated from soil from Dokdo, Korea. *Antonie van Leeuwenhoek*, **109**, 1397-1402, doi: 10.1007/s10482-016-0738-x.

Kim, J. J., Masui, R., Kuramitsu, S., Seo, J. H., Kim, K. & Sung, M. H. 2008. Characterization of growth-supporting factors produced by *Geobacillus toebii* for the commensal thermophile *Symbiobacterium toebii*. *Journal of Microbiology and Biotechnology*, **18**, 490-496.

Kim, K., Kim, J. J., Masui, R., Kuramitsu, S. & Sung, M. H. 2011. A commensal symbiotic interrelationship for the growth of *Symbiobacterium toebii* with its partner bacterium, *Geobacillus toebii*. *BMC Research Notes*, 4, doi: 10.1186/1756-0500-4-437.

Kim, S. -J., Ahn, J. -H., Lee, T. -H., Weon, H. -Y., Hong, S. -B., Seok, S. -J., Whang, K. -S. & Kwon, S. -W. 2013. *Reyranella soli* sp. nov., isolated from forest soil, and emended description of the genus *Reyranella* Pagnier et al. 2011. *International Journal of Systematic and Evolutionary Microbiology*, **63**, 3164-3167, doi: 10.1099/ijs.0.045922-0.

Kobayashi, K. & Aoyagi, H. 2019. Microbial community structure analysis in *Acer palmatum* bark and isolation of novel bacteria IAD-21 of the phylum Abditibacteriota (former candidate division FBP). *PeerJ*, **7**, doi: 10.7717/peerj.7876.

Lai, Q., Yuan, J., Wu, C. & Shao, Z. 2009. *Oceanibaculum indicum* gen. nov., sp. nov., isolated from deep seawater of the Indian Ocean. *International Journal of Systematic and Evolutionary Microbiology*, **59**, 1733-1737, doi: 10.1099/ijs.0.004341-0.

Lee, H., Kim, D. -U., Lee, S., Park, S., Yoon, J. -H., Seong, C. N. & Ka, J. -O. 2017. *Reyranella terrae* sp. nov., isolated from an agricultural soil, and emended description of the genus *Reyranella*. *International Journal of Systematic and Evolutionary Microbiology*, **67**, 2031-2035, doi: 10.1099/ijsem.0.001913.

Lee, J. -C. & Whang, K. -S. 2014. *Reyranella graminifolii* sp. nov., isolated from bamboo (*Phyllostachys bambusoides*) litter. *International Journal of Systematic and Evolutionary Microbiology*, **64**, 2503-2507, doi: 10.1099/ijs.0.062968-0.

Lennon, J. & Jones, S. 2011. Microbial seed banks: the ecological and evolutionary implications of dormancy. *Nature Reviews Microbiology*, **9**, 119–130, doi: 10.1038/nrmicro2504.

Liu, Y., Jin, J. -H., Liu, Y. -H., Zhou, Y. -G. & Liu, Z. -P. 2010. *Dongia mobilis* gen. nov., sp. nov, a new member of the family *Rhodospirillaceae* isolated from a sequencing batch reactor for treatment of malachite green effluent. *International Journal of Systematic and Evolutionary Microbiology*, **60**, 2780-2785, doi: 10.1099/ijs.0.020347-0.

Martin, V. N. R., Schwab, C., Krych, L., Voney, E., Geirnaert, A., Braegger, C. & Lacroix, C. 2019. Colonization of *Cutibacterium avidum* during infant gut microbiota establishment. *FEMS Microbiology Ecology*, **95**, doi: 10.1093/femsec/fiy215.

Maynaud, G., Willems, A., Soussou, S., Vidal, C., Mauré, L., Moulin, L., Cleyet-Marel, J. -C. & Brunel, B. 2012. Molecular and phenotypic characterization of strains nodulating *Anthyllis vulneraria* in mine tailings, and proposal of *Aminobacter anthyllidis* sp. nov., the first definition of *Aminobacter* as legume-nodulating bacteria. *Systematic and Applied Microbiology*, **35**, 65-72, doi: 10.1016/j.syapm.2011.11.002.

Noviana, Z., Vieira, S., Pascual, J., Fobofou, S. A. T., Rohde, M., Spröer, C., Bunk, B. & Overmann, J. 2020. *Hypericibacter terrae* gen. nov., sp. nov. and *Hypericibacter adhaerens* sp. nov., two new members of the family *Rhodospirillaceae* isolated from the rhizosphere of *Hypericum perforatum*. *International Journal of Systematic and Evolutionary Microbiology,* **70**, 1850-1860, doi: 10.1099/ijsem.0.003983.

Overmann, J. 2013. Chapter 7. Principles of enrichment, isolation, cultivation, and preservation of prokaryotes. *In: The Prokaryotes: Prokaryotic Biology and Symbiotic Associations*. 149–207, doi: 10.1007/978-3642-30194-0_7.

Pagnier, I., Raoult, D. & La Scola, B. 2011. Isolation and characterization of *Reyranella massiliensis* gen. nov., sp. nov. from freshwater samples by using an amoeba co-culture procedure. *International Journal of Systematic and Evolutionary Microbiology*, **61**, 2151-2154, doi: 10.1099/ijs.0.025775-0.

Peng, J., Hao, B., Liu, L., Wang, S., Ma, B., Yang, Y., Xie, F. & Li, Y. 2014. RNA-Seq and Microarrays Analyses Reveal Global Differential Transcriptomes of *Mesorhizobium huakuii* 7653R between Bacteroids and Free-Living Cells. *PLoS ONE*, **9**, doi: 10.1371/journal.pone.0093626.

Pester, M., Bittner, N., Deevong, P., Wagner, M. & Loy, A. 2010. A 'rare biosphere' microorganism contributes to sulfate reduction in a peatland. *The ISME Journal*, **4**, 1591-1602 (2010), doi: 10.1038/ismej.2010.75.

Rahalkar, M., Bahulikar, R. A., Deutzmann, J. S., Kroth, P. G. & Schink, B. 2012. *Elstera litoralis* gen. nov., sp. nov., isolated from stone biofilms of Lake Constance, Germany. *International Journal of Systematic and Evolutionary Microbiology*, **62**, 1750-1754, doi: 10.1099/ijs.0.026609-0.

Richter, M. & Roselló-Móra, R. 2009. Shifting the genomic gold standard for the prokaryotic species definition. *Proceedings of the National Academy of Sciences*, **106**, 19126-19131, doi: 10.1073/pnas.0906412106.

Roesch, L. F. W., Fulthorpe, R. R., Riva, A., Casella, G., Hadwin, A. K. M., Kent, A. D., Daroub, S. H., Camargo, F. A. O., Farmerie, W. G. & Triplett, E. W. 2007. Pyrosequencing enumerates and contrasts soil microbial diversity. The ISME Journal, **1**, 283-290, doi: 10.1038/ismej.2007.53.

Sandaa, R. -A., Torsvik, V., Enger, Ø., Daae, F. L., Castberg, T. & Hahn, D. 1999. Analysis of bacterial communities in heavy metal-contaminated soils at different levels of resolution. FEMS Microbiology Ecology, 30, 237-251, doi: 10.1111/j.1574-6941.1999.tb00652.x.

Sansinenea, E. & Ortiz, A. 2011. Secondary metabolites of soil *Bacillus* spp. *Biotechnology Letters*, **33**, 1523-1538, doi: 10.1007/s10529-011-0617-5.

Shalem Raj, P., Kalyana Chakravarthy, S., Ramaprasad, E. V. V., Sasikala, Ch. & Ramana, ChV. 2012 *Phaeospirillum tilakii* sp. nov., a phototrophic alphaproteobacterium isolated from aquatic sediments. *International Journal of Systematic and Evolutionary Microbiology*, 62, 1069-1074, doi: 10.1099/ijs.0.032250-0.

Sheu, S. -Y., Chen, Y. -L, Young, C. -C. & Chen, W. -M. 2013. *Lacibacterium aquatile* gen. nov., sp. nov., a new member of the family *Rhodospirillaceae* isolated from a freshwater lake. *International Journal of Systematic and Evolutionary Microbiology*, **63**: 4797-4804, doi: 10.1099/ijs.0.055145-0.

Stewart, E. J. 2012. Growing Unculturable Bacteria. *Journal of Bacteriology*, **194**, 4151-4160, doi: 10.1128/JB.00345-12.

Su, Y., Wang, R., Sun, C., Han, S., Hu, J., Wu, D., Ma, Z., Chen, J. & Wu, M. 2016. *Thalassobaculum fulvum* sp. nov., isolated from deep seawater. *International Journal of Systematic and Evolutionary Microbiology*, **66**, 2186-2191, doi: 10.1099/ijsem.0.001008.

Tanaka, Y., Hanada, S., Manome, A., Tsuchida, T., Kurane, R., Nakamura, K. & Kamagata, Y. 2004. *Catellibacterium nectariphilum* gen. nov., sp. nov., which requires a diffusible compound from a strain related to the genus Sphingomonas for vigorous growth. *International Journal of Systematic and Evolutionary Microbiology*, **54**, 955-959, doi:10.1099/ijs.0.02750-0.

Tiirola, M. A., Männistö, M. K., Puhakka, J. A. & Kulomaa, M. S. 2002. Isolation and characterization of *Novosphingobium* sp. strain MT1, a dominant polychlorophenol-degrading strain in a groundwater bioremediation system. *Applied and Environmental Microbiology*, **68**, 173-180, doi: 10.1128/AEM.68.1.173-180.2002.

Urios, L., Michotey, V., Intertaglia, L., Lesongeur, F. & Lebaron, P. 2008. *Nisaea denitrificans* gen. nov., sp. nov. and *Nisaea nitritireducens* sp. nov., two novel members of the class Alphaproteobacteria from the Mediterranean Sea. *International Journal of Systematic and Evolutionary Microbiology*, **58**, 2336-2341, doi: 10.1099/ijs.0.64592-0.

Urios, L., Michotey, V., Intertaglia, L., Lesongeur, F. & Lebaron, P. 2010. *Thalassobaculum salexigens* sp. nov., a new member of the family *Rhodospirillaceae* from the NW Mediterranean Sea, and emended description of the genus *Thalassobaculum*. *International Journal of Systematic and Evolutionary Microbiology*, **60**, 209-213, doi: 10.1099/ijs.0.011460-0.

Vasilyeva, L. V. 1985. *Stella*, a new genus of soil prosthecobacteria, with proposals for *Stella humosa* sp. nov. and *Stella vacuolata* sp. nov *International Journal of Systematic and Evolutionary Microbiology*, **35**, 518-521, doi: 10.1099/00207713-35-4-518.

Yamada, K., Fukuda, W., Kondo, Y., Miyoshi, Y., Atomi, H. & Imanaka, T. 2011. *Constrictibacter antarcticus* gen. nov., sp. nov., a cryptoendolithic micro-organism from Antarctic white rock. *International Journal of Systematic and Evolutionary Microbiology*, **61**, 1973-1980, doi: 10.1099/ijs.0.026625-0.

Yoon, S. -H., Ha, S. -M., Kwon, S., Lim, J., Kim, Y., Seo, H. & Chun, J. 2017a. Introducing EzBioCloud: A taxonomically united database of 16S rRNA and whole genome assemblies. *International Journal of Systematic and Evolutionary Microbiology*, 2017a, **67**, 1613-1617, doi: 10.1099/ijsem.0.001755.

Yoon, S. -H., Ha, S. -M., Lim, J., Kwon, S. & Chun, J. 2017b. A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie van Leeuwenhoek*, **110**, 1281–1286, doi: 10.1007/s10482-017-0844-4.

Zhang, D., Yang, H., Zhang, W., Huang, Z. & Liu, S. -J. 2003. *Rhodocista pekingensis* sp. nov., a cyst-formingphototrophic bacterium from a municipal wastewater treatment plant. *International Journal of Systematic and Evolutionary Microbiology*, **53**, 1111-1114, doi: 10.1099/ijs.0.02500-0.

Zhang, G. I., Hwang, C. Y. & Cho, B. C. 2008. *Thalassobaculum litoreum* gen. nov., sp. nov., a member of the family *Rhodospirillaceae* isolated from coastal seawater. *International Journal of Systematic and Evolutionary Microbiology*, **58**, 479-485, doi: 10.1099/ijs.0.65344-0.

Chapter 6 – Conclusion

The plant genus *Hypericum* is widely known for the production of hypericin and hyperforin with anti-anxiety and anti-depression effect. Recent studies also suggest neuroprotective traits against Alzheimer disease. Studies of *Hypericum* plants are mostly focused on the secondary metabolites while the associated bacterial community is rarely investigated albeit the potential in modulating the plant secondary metabolites. Therefore, this study focused on the bacterial communities in the rhizosphere and roots of *Hypericum* plants to gain better understanding of the potential link between the associated bacterial communities and hypericin and hyperforin production in the host plant.

Hypericum bacterial community is governed by habitat type (bulk soil, rhizosphere and inside the roots), followed by soil substrate (with distinct pH) and plant species. Investigation of each specific habitat revealed that plant species strongly drives the rhizosphere and root bacterial communities. The structure and composition of the total (DNA-based) and active (RNA-based) bacterial communities are different, implying differences on the transcriptional activities across resident bacteria. We further identify potentially active taxa based on the values of rRNA:rDNA ratio that might trigger hypericin and hyperforin. These taxa are detected to be active in the rhizosphere or roots of hypericin and hyperforin-producing species, but not in non-producer. However, a more comprehensive study is needed to confirm the result and to reveal specific ecological functions that these active taxa carry out in the rhizosphere or roots of the host plant.

Furthermore, the employment of growth media with low-nutrient content combined with dilution-to-extinction approach and a high-throughput screening (by sequencing) is promising to improve cultivation success. In addition, co-cultivation is crucial for isolating the unculturable taxa with missing biological pathways. These strategies should be employed in the future to isolate novel bacterial taxa that may elicit hypericin and hyperforin in *Hypericum*.

This study aims to provide insights into bacterial community in the rhizosphere and roots of *Hypericum* plants, as a step closer for a better understanding of the link between the bacterial communities and the hypericin and hyperforin production in the host plant. However, the study is still at its early phase and a deeper study is needed to validate the here presented results, especially of the active taxa in the rhizosphere and roots of hypericin and hyperforin-producing species that potentially trigger the compounds. Since many of these active taxa are unculturable, further studies need to be accompanied with a premeditated cultivation effort.

Acknowledgements

I would like to express my sincere gratitude to the following people without whom this dissertation would not have been possible.

First and foremost, I would like to express my deepest appreciation to Prof. Jörg Overmann for the opportunity to do my studies in the microbial ecology and diversity group of DSMZ, with a very interesting topic regarding the bacterial community of the plant genus *Hypericum.* The research is only possible with his expertise, support, and guidance.

For the great guidance in the data analysis, thesis writing, and preparation of several scientific talks, along with many fruitful discussions on the research topic, I sincerely thank Johannes Sikorski.

For the great assistance in reviewing my scientific paper and many helpful discussions regarding the laboratory work, I would like to thank Selma Vieira and Javier Pascual.

I would like to express my deepest gratitude towards Alicia Geppert, Anika Methner, and Franziska Burkart for the great assistance in the laboratory work, especially regarding the harvest of *Hypericum* plants and the processing of hundreds of samples prior to the bacterial community analysis.

I would like to offer my special thanks to my colleagues in DSMZ, for the enjoyable working environment and for sharing extensive knowledge in the topic of microbiology. I learnt a lot during my stay at the institute.

I am deeply grateful to my colleagues at Leibniz IPB Halle and Botanic Garden and Botanical Museum Berlin-Dahlem for providing me with the *Hypericum* plants. This work would not have been possible without their helping hands.

For my dearest friends, Senlie Octaviana, Gian Primahana, Mardy Zaenal, Ingrid Meyer, Faiza Maryani, and Chandra Risdian, I am very thankful for the support, motivation, and all the nice memories I had throughout this journey.

For the nice hiking trips in the Hartz mountain during my study and the weekly badminton session, I would like to thank Klaus-Dieter Aumann, Ulrike Beutling, and Enge Sudarman.

My deepest gratitude goes to my family, especially my beloved parents, Dede Margo Irianto and Yani Yusnia, my older and younger brothers, Adhi Nugraha Putra Perdana and Rizky Akbar, and my

sister-in-law, Ima Humaeroh, for their continuous support and long-lasting motivation. They keep me going throughout this very challenging journey.

Last but certainly not least, I am deeply grateful for the Deutscher Akademischer Austauschdienst (DAAD) for the financial support that I had received during my studies. This work would not have been possible without their support.

Curriculum vitae

Personal Data

Name	:	Zahra Noviana
Nationality	:	Indonesian
Place of Birth	:	Bandung, Indonesia
Date of Birth	:	7th November 1990

Formal Education

October 2016 – November 2020	**Doctoral studies** Microbial ecology and diversity (MÖD) Leibniz Institute DSMZ – German Collection of Microorganisms and Cell Cultures
June 2012 – June 2013	**Master degree** Study program of biotechnology School of life sciences and technology Institut Teknologi Bandung
June 2008 – June 2012	**Bachelor degree** Study program of microbiology School of life sciences and technology Institut Teknologi Bandung

Academic Projects

October 2016 – November 2020	**PhD dissertation** *Structure and composition of bacterial communities in the rhizosphere of Hypericum species* Leibniz Institute DSMZ – German Collection of Microorganisms and Cell Cultures (Prof. Dr. Jörg Overmann)
2015	**SATREPS Research Subject 4** *Study of chicken gut microbiota by T-RFLP* Research Center of Biology – Indonesian Institute of Sciences & Japan Collection of Microorganisms (JCM) – RIKEN BioResource Research Center
September 2012 – May 2013	**Master degree dissertation** *Optimization of tofu fermentation through variation of percentages and ratio of inoculums to increase protein and isoflavone content* Study program of biotechnology, School of life sciences and technology, Institut Teknologi Bandung (Dr. Dea Indriani Astuti)

December 2011 – May 2012 **Bachelor degree dissertation**
Optimization of tofu fermentation by variation of the ratio of inoculums which were isolated from tofu starter
Study program of microbiology, School of life sciences and technology, Institut Teknologi Bandung (Dr. Dea Indriani Astuti)

Publications

Noviana, Z., Vieira, S., Pascual, J., Fobofou, S. A. T., Rohde, M., Spröer, C., Bunk, B. & Overmann, J. 2020. *Hypericibacter terrae* gen. nov., sp. nov. and *Hypericibacter adhaerens* sp. nov., two new members of the family *Rhodospirillaceae* isolated from the rhizosphere of *Hypericum perforatum*. *International Journal of Systematic and Evolutionary Microbiology,* 70: 1850-1860.

Astuti, D. I., & **Noviana, Z.** 2013. Optimization of fermented tofu with high isoflavone content through variation of inoculum percentages and ratios of *Lactobacillus plantarum, Lactobacillus acidophilus, and Leuconostoc mesenteroides*. ***Journal of Mathematical and Fundamental Sciences***, 45: 263-273.

Posters

Noviana, Z., Sikorski, J., Overmann, J., 2020. Structure and composition of bacterial communities in the rhizosphere of *Hypericum* species. Vereinigung für Allgemeine und Angewandte Mikrobiologie Jahrestagung 2020, Leipzig.